



Published in final edited form as:

Neuroimage. 2021 January 01; 224: 117399. doi:10.1016/j.neuroimage.2020.117399.

Approximating anatomically-guided PET reconstruction in image space using a convolutional neural network

Georg Schramm^{a,*}, David Rigie^b, Thomas Vahle^c, Ahmadreza Rezaei^a, Koen Van Laere^a, Timothy Shepherd^d, Johan Nuyts^a, Fernando Boada^b

^aDepartment of Imaging and Pathology, Division of Nuclear Medicine, KU/UZ Leuven, Leuven, Belgium

^bCenter for Advanced Imaging Innovation and Research (CAI2R) and Bernard and Irene Schwartz Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine, NYC, US

^cSiemens Healthcare GmbH, Erlangen, Germany

^dDepartment of Neuroradiology, NYU Langone Health, Department of Radiology, New York University School of Medicine, New York, US

Abstract

In the last two decades, it has been shown that anatomically-guided PET reconstruction can lead to improved bias-noise characteristics in brain PET imaging. However, despite promising results in simulations and first studies, anatomically-guided PET reconstructions are not yet available for use in routine clinical because of several reasons. In light of this, we investigate whether the improvements of anatomically-guided PET reconstruction methods can be achieved entirely in the image domain with a convolutional neural network (CNN). An entirely image-based CNN post-reconstruction approach has the advantage that no access to PET raw data is needed and, moreover, that the prediction times of trained CNNs are extremely fast on state of the art GPUs which will substantially facilitate the evaluation, fine-tuning and application of anatomically-guided PET reconstruction in real-world clinical settings.

In this work, we demonstrate that anatomically-guided PET reconstruction using the asymmetric Bowsher prior can be well-approximated by a purely shift-invariant convolutional neural network in image space allowing the generation of anatomically-guided PET images in almost real-time.

We show that by applying dedicated data augmentation techniques in the training phase, in which 16 [¹⁸F]FDG and 10 [¹⁸F]PE2I data sets were used, lead to a CNN that is robust against the used

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

*Corresponding author, georg.schramm@kuleuven.be (G. Schramm).

Credit authorship contribution statement

Georg Schramm: Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Visualization. **David Rigie:** Methodology, Software. **Thomas Vahle:** Methodology, Software, Data curation. **Ahadreza Rezaei:** Methodology, Software. **Koen Van Laere:** Resources, Writing - review & editing, Funding acquisition. **Timothy Shepherd:** Resources, Data curation, Writing - review & editing. **Johan Nuyts:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition, Project administration. **Fernando Boada:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition, Project administration.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.neuroimage.2020.117399](https://doi.org/10.1016/j.neuroimage.2020.117399).

PET tracer, the noise level of the input PET images and the input MRI contrast. A detailed analysis of our CNN in 36 [^{18}F]FDG, 18 [^{18}F]PE2I, and 7 [^{18}F]FET test data sets demonstrates that the image quality of our trained CNN is very close to the one of the target reconstructions in terms of regional mean recovery and regional structural similarity.

Keywords

Molecular imaging; Magnetic resonance imaging; Image reconstruction; Quantification; Machine learning

1. Introduction

Positron emission tomography (PET) is an important quantitative clinical molecular imaging technique with high sensitivity and high specificity that is increasingly used in oncology, neurology, infection/inflammation, and cardiology. However, compared with structural imaging modalities such as x-ray computer tomography (CT) or magnetic resonance imaging (MRI), it suffers from a number of limitations. First, due to physical and technical constraints, acquired PET data are contaminated by a high level of Poisson noise which propagates into the reconstructed images. Second, reconstructed PET images from clinical whole-body scanners have a relatively low spatial resolution of ca. 4–5 mm. Such relative low resolution and the resulting partial volume effects (Erlandsson et al., 2012; Fazio and Perani, 2000; Hutton et al., 2013; Mueller-Gaertner et al., 1992; Yang et al., 1996) hinder accurate quantification of tracer uptake within small structures, such as small lymph nodi and grey matter nuclei or atrophic cortex, leading to a concomitant decrease in sensitivity and specificity. Moreover, it reduces the detection of small structures with low contrast.

To overcome these challenges, especially in PET brain imaging, many techniques for anatomically-guided PET reconstruction using high-resolution structural MR images have been developed in the past (see e.g. Baete et al., 2004; Bland et al., 2017; Bowsher et al., 2004; Comtat et al., 2002; Ehrhardt et al., 2016; Knoll et al., 2016; Lipinski et al., 1997; Loeb et al., 2015; Mehranian et al., 2018; Nuyts, 2007; Schramm et al., 2017; Vunckx et al., 2012; Vunckx et al., 2014; Vunckx and Nuyts, 2010). These techniques have received increased interest recently due to the introduction of whole-body, clinical PET/MRI scanners during the last decade, which effectively removed all logistical and practical constraints for procuring the required high-resolution MR images.

Unfortunately up to now, despite promising results in simulations and first studies (Shepherd et al., 2019), anatomically-guided PET reconstructions are not available for routine clinical use. The reasons for that are manifold. First, until now, no vendor has implemented any of the anatomically-guided PET reconstructions into their products which prohibits the use and evaluation by standard clinical users on large patient populations. Furthermore, these methods are not applied even in the few academic centers that are able to perform anatomically-guided PET reconstructions. This is because anatomically-guided PET reconstructions are usually rather time-consuming (ca. 5 min - 1.5 h depending on the used hardware and the data size) which complicates the inevitable (task-based) tuning of the prior weight (β) that controls the degree of anatomically-guided regularization. Moreover, so far,

validation of anatomically-guided PET reconstructions in big cohorts for specific clinical tasks is missing which limits the trust of clinicians in those reconstructions. The latter is especially true in the case where structures such as gradients are not shared between the PET and MR images.

In light of this, we sought to investigate whether the improvements of anatomically-guided PET reconstruction methods can be achieved entirely in the image domain with a convolutional neural network (CNN). The purpose of this work is to train a CNN to transform standard clinical PET and high-resolution MR images into an anatomically-guided PET reconstruction. Such image-based CNNs would have the advantage that they could be easily deployed due to the availability of excellent free and open-source toolboxes (Abadi et al., 2015; Chollet, 2015). They could also be applied retrospectively to all PET/MRI studies available in clinical picture archiving systems, even if the PET raw data is no longer available. Finally, since the predictions made by CNNs are very fast (a prediction of our network on a recent GPU takes ca. 1–2 s), the use of an image-based CNN approach dramatically facilitates the application, fine-tuning (by using multiple CNNs trained with different target β s), and evaluation of anatomically-guided PET reconstructions for a wider range of clinical applications.

1.1. Related work

Deep learning techniques using convolutional neural networks have shown great potential in different classical computer vision tasks such as denoising and image upsampling (super-resolution problem) as e.g. demonstrated in Dong et al. (2016); Liu et al. (2018); Zhang et al. (2018). The use of convolutional neural networks in the field of PET image post-processing and PET image reconstruction has also strongly increased in the last years. Most published works focus on convolutional neural networks designed and trained to predict high count from low count input PET images to reduce the injected activity or the acquisition time. Cui et al. (2017) proposed a stacked sparse auto-encoder framework to denoise dynamic PET images. Xiang et al. (2017) designed a 2D convolutional neural network that predicts a high count PET image from an input low count PET image and an aligned T1-weighted MR image. Wang et al. (2018) proposed a 3D conditional generative adversarial networks to predict high count PET images from low-count ones. Chen et al. (2019), designed a U-net like convolutional layer to predict full-dose PET images from ultra low dose PET images (1% of the full dose) and an aligned T1 MR image. In a more advanced approach, Kim et al. (2018) proposed an iterative reconstruction approach using a deep learned prior to obtain low noise PET reconstructions under variable noise conditions. Yang et al. (2018) designed a multilayer perceptron neural network to predict enhanced PET image patches below the classically achievable bias-variance trade-off line from a series of reconstructed patches with different degrees of regularization using the Green smoothness (log cosh) prior. Gong et al. (2019b) showed that their proposed approach of embedding a pre-trained deep residual convolutional neural network exploiting inter-patient information into iterative PET reconstruction is superior to the post-processing approach using a convolutional neural network for denoising. Moreover, the same authors also showed in Gong et al. (2019a) that a technique similar to the deep image prior method (Ulyanov et al., 2018) can be used to improve iterative anatomically-guided PET reconstruction without the

need for training pairs. Hong et al. (2018) proposed a deep residual convolutional neural network to perform super-resolution on PET sinograms before image reconstruction to improve the reconstructed spatial resolution. Häggström et al. (2019) introduced a deep convolutional encoder-decoder network called DeepPET that was trained to directly predict 2D PET reconstructions from pre-corrected 2D non-time-of-flight sinogram data.

1.2. Aim of this work

In contrast to all the works of the previous subsection, the aim of this work is different. Here, we aimed to design and train a purely image-based convolutional neural network that learns to predict anatomically-guided PET reconstructions in almost real-time using standard clinical OSEM PET images and high-resolution structural MR images as input. Recently (Rigie et al., 2018), we demonstrated in an initial proof of concept study that this is possible for data from a single scanner and a single PET tracer and similar noise levels. In this work, we demonstrate that this approach can be extended to the use of multiple tracers, multiple PET scanners and noise levels and that the resulting CNN is very robust and extensible to other tracers without the need of retraining.

2. Materials and methods

2.1. Anatomically-guided PET reconstruction

Since the aim of this work is to approximate anatomically-guided PET reconstruction with a convolutional neural network in image space, we first give a brief overview of the former. This description is based on our earlier work in Schramm et al. (2017).

In general, the penalized-likelihood PET reconstruction problem using an MRI-based penalty function can be written as

$$\arg \min_{\mathbf{u} \geq 0} \underbrace{\sum_i \hat{y}_i(\mathbf{u}) - y_i \log \hat{y}_i(\mathbf{u})}_{L_{\text{PET}}(\hat{\mathbf{y}}(\mathbf{u}))} + \beta \mathbf{R}(\mathbf{u}, \mathbf{v}), \quad (1)$$

where y_i are measured coincidences for line of response i ¹, L_{PET} is the negative of the Poisson log-likelihood, \mathbf{u} is the non-negative discretized PET image to be reconstructed, \mathbf{R} is a penalty function (prior) and β is a non-negative scalar controlling the weight of the regularization. The forward model $\mathbf{u} \mapsto \hat{\mathbf{y}}(\mathbf{u})$ for computing the expected coincidences $\hat{\mathbf{y}}$ is given by

$$\hat{\mathbf{y}}(\mathbf{u}) = \mathbf{P}\mathbf{u} + \mathbf{s}. \quad (2)$$

In (2), the operator \mathbf{P} with matrix elements P_{ij} is the forward projection including the effects of sensitivity, attenuation, finite spatial resolution, and time of flight (if available). The noise-free estimate of additive contaminations such as random and scattered events is denoted as \mathbf{s} . In the case of anatomically-guided PET reconstruction, \mathbf{R} is a function that incorporates information from an aligned anatomical prior MR image \mathbf{v} .

¹In the case of time-of-flight (TOF) PET reconstruction, i is a master index that includes the physical lines of response and the TOF bins.

To investigate whether learning anatomically-guided PET reconstruction is possible, we focus in this proof of concept work on one specific prior function, namely the asymmetric Bowsher prior (Bowsher et al., 2004; Schramm et al., 2017; Vunckx and Nuyts, 2010). This prior, proposed in Vunckx and Nuyts (2010) as a refinement of Bowsher et al. (2004), is a smoothing Markov prior operating on a position-dependent set of voxels. In the discrete setting, it can be written as

$$R_{\text{BOW}}(\mathbf{u}, \mathbf{v}) = \sum_j \sum_k w_{jk}(\mathbf{v}) \mathbf{M}(u_j, u_k), \quad (3)$$

where $\mathbf{M}(u_j, u_k)$ is a function penalizing differences between the reconstructed PET uptake in voxels k and j . The neighbor weights w_{jk} are computed from the aligned MR image. w_{jk} is 1 if voxel k is in a set B_j , which consists of the n most similar voxels around j , and 0 otherwise. In all reconstructions shown in this work, this set B_j includes the 4 out of the 18 nearest voxels around j that, according to the MR image intensity v , are most similar to voxel j . The absolute MRI intensity difference was used as the similarity measure. For \mathbf{M} the relative difference

$$\mathbf{M}(a, b) = \frac{(a - b)^2}{a + b}, \quad (4)$$

as defined in Nuyts et al. (2002) with $\gamma = 0$ was used. Solving the penalized PET reconstruction problem (1) is commonly done using iterative algorithms. In this work, the heuristic algorithm proposed in Nuyts et al. (2002) and applied to PET reconstruction with the asymmetric Bowsher prior using the relative difference prior in Vunckx and Nuyts (2010) was used for all reconstructions. The computation time of those iterative algorithms can be high since usually several iterations (updates) containing many time-consuming forward and back projection operations are needed to reach convergence. Depending on the computational efficiency of the forward and back projectors, the computational hardware, the optimization algorithm, the voxel size, and the size of the (TOF) projection data, solving (1) can take between several minutes and several hours.

2.2. Data sets

The number of data sets that were used in the work to train, validate and test our convolutional neural network are listed in detail in Table 1. For training and validation (monitoring of potential over-fitting and adaptation of the learning rate), 26 and 6 subject acquisitions with two different tracers ($[^{18}\text{F}]\text{FDG}$ - tracer for glucose metabolism, $[^{18}\text{F}]\text{PE2I}$ - tracer for dopamine transporter) from two different PET/MRI scanners - the Biograph mMR (Siemens Healthcare, Erlangen, Germany) (Delso et al., 2011) and the GE SIGNA PET/MR (GE Healthcare, Milwaukee, US) (Grant et al., 2016) - were used, respectively. For the mMR $[^{18}\text{F}]\text{FDG}$ acquisitions, 1 min, and 3 min were created in addition to the 20 min data sets to generate $[^{18}\text{F}]\text{FDG}$ OSEM input images with different noise levels. Note that in the training phase of the CNN, the Bowsher reconstructions derived from the full 20 min data were always used as target which means that our network is also implicitly trained to perform a “low counts to high counts” mapping. Evaluation (testing) of the trained network was performed on 54 cases acquired with the two tracers

used for training and on 7 additional cases acquired with the [^{18}F]FET (tracer for amino acid metabolism for brain tumors) that were not used during training.

All data were retrospectively collected with approval from the NYU institutional review board and the Ethical Committee of UZ Leuven, respectively.

2.3. Data acquisition and image reconstruction

Before training, validation, and testing of our convolutional neural network, conventional ordered subset maximum expectation maximization (OSEM) PET images and anatomically-guided PET images using the asymmetric Bowsher prior with a prior weight of $\beta = 10$ were reconstructed using in-house developed prototype reconstruction software. The prior weight was chosen in a previous study based on FDG brain images by an experienced radiologist. Note that $\beta = 10$ is not necessarily the optimal regularization weight for non-FDG tracers depending on the clinical task. However, careful tuning of the regularization weight β is beyond the scope of this work, and, moreover, network training can be repeated with target reconstructions that used a different level of regularization.

All PET reconstruction parameters are listed in Table 2. The forward and back projection were implemented using the distance driven approach (De Man and Basu, 2004). For the OSEM images, no post smoothing was applied.

For the anatomically-guided PET reconstructions, MPRAGE (sequence parameters: imaging mode 3D, flip angle 12° , TI 900 ms, TE 2.75 ms, TR 2400 ms, voxel size $1 \times 1 \times 1 \text{ mm}^3$) and BRAVO (sequence parameters: imaging mode 3D, flip angle 12° , TI 450 ms, voxel size $0.7 \times 1 \times 1 \text{ mm}^3$, TE 3.2 ms, TR 9.5 ms) MRI acquisitions were used as prior images for the mMR and SIGNA data sets, respectively. To mitigate the influence of small residual misalignments, e.g. due to patient motion, the anatomical prior MR image was always rigidly coregistered to the OSEM PET reconstruction by maximizing mutual information using in-house developed software.

2.4. Convolutional neural network design

Fig. 1 shows the network architecture chosen to learn anatomically-guided PET reconstruction in image space. The network takes 2 3D images (OSEM PET and structural MRI) of size (n_1, n_2, n_3) as input and outputs an anatomically-guided PET reconstruction of the same size.

The network consists of 8 3D convolutional layers using $3 \times 3 \times 3$ kernels and 30 features followed by a parametric rectifying unit (PreLU). The features of the first convolutional layer are split into 15 PET and 15 MRI only features which are concatenated before the second convolutional layer. A final 3D convolutional layer using a $1 \times 1 \times 1$ kernel and a PreLU was added to guarantee that the size of the 3D output image is the same as the size of the input. After the last convolutional layer, the PET input is added to the output such that the network learns to predict the residual between the OSEM input PET and the anatomically-guided target PET image. Note that since the spatial size of the output of every layer is kept constant, which is achieved by using padding with zeros before applying the convolutions, 3D images of arbitrary size can be given as input to the network. This, in turn,

allows that the network can be trained on small patches, but the predictions for the test cases can be done on the whole images provided that enough memory is available.

Transforming an OSEM PET image with the help of a structural MR image into an anatomically-guided PET image is related to a constrained deblurring and denoising problem. The latter is a local problem and shift-invariant if we assume that the OSEM PET point spread function is shift-invariant, which holds for the case of brain PET imaging with a whole-body PET system. Hence, we decided to use a purely convolutional shift-invariant neural network without any fully connected, downsampling or upsampling layers.

The fact that our problem is local, and the fact that the network is shift invariant makes it possible to train the network on relatively small 3D patches. This has the advantage that, first, the training becomes much faster, and second that we can create many small training patches from every single PET/MRI acquisition. This, in turn, allows to train the network on a relatively small number of PET/MRI acquisitions.

In total, our network has 171421 free parameters that were optimized during training.

2.5. Data preprocessing, augmentation, and CNN training

To train our convolutional neural network, pairs of input OSEM PET and structural MR images and target anatomically-guided PET reconstructions using the asymmetric Bowsheer prior (BOW) reconstructed with a prior weight of $\beta = 10$ were used. Before training and prediction, all OSEM PET and MR input images were trilinearly interpolated to a common voxel size of $1 \times 1 \times 1 \text{ mm}^3$ and normalized by dividing each image by its 99.9% percentile.² Moreover, the field of view was cropped to the bounding box of the head in the MR image. For training, the target BOW images were interpolated and cropped in the same way and the normalization factor of the input OSEM images were applied. Data augmentation was performed by randomly flipping the contrast of the input MR image, randomly applying spatial flipping in all 3 directions of input and target images, and by using [¹⁸F]FDG OSEM images with different acquisition times and thus different noise levels. All those data augmentation techniques were done to obtain a trained network that is robust against the input MRI contrast, the patient anatomy, and the noise level of the input OSEM PET image. Training was performed in batches of 64 small 3D patches of size (29,29,29). The data set and the location of every 3D patch was randomly sampled in all batches. The mean square error loss function was optimized using the Adam optimizer with an initial learning rate of 10^{-3} and 10,000 epochs with 20 steps per epoch resulting in 200,000 updates of the network weights. Note that since the patch extraction is completely randomized, the definition of an epoch is arbitrary. The learning rate was reduced by a factor of 2 as soon as the validation loss reached a plateau, but limited to be at least 10^{-4} . The network weights were initialized with random numbers drawn from a normal distribution with standard deviation of 0.07. Training of the network was implemented in Keras v2.2.2 (Chollet, 2015) with tensorflow v1.9 (Abadi et al., 2015) backend and performed on four Nvidia P100 GPUs resulting in a

²The output of our CNN is scaled with the inverse of the individual input normalization factor of the OSEM PET image to preserve absolute quantification.

training time of approximately 8.5 h. Supplementary Fig. S9 shows the training and validation loss as a function of the epoch.

2.6. Quantitative regional analysis

The quality of the anatomically-guided PET images predicted by the CNN (BOW_{CNN}) with respect to the ground truth of the iteratively reconstructed anatomically-guided PET images (BOW) was evaluated based on two metrics. First, we computed regional recovery coefficients

$$RC_{\text{mean}}(\text{ROI}) = \frac{ROI_{\text{mean}}(BOW_{CNN})}{ROI_{\text{mean}}(BOW)}, \quad (5)$$

in different regions of interests (ROIs) defined by Freesurfer v6 (Fischl et al., 2004) segmentations of the structural MR images, where $ROI_{\text{mean}}(x)$ denotes the mean signal intensity of image x in a given ROI. Second, we calculated the average regional structural similarity as:

$$SSIM_{\text{mean}}(\text{ROI}) = ROI_{\text{mean}}(SSIM(BOW_{CNN}, BOW)), \quad (6)$$

where $SSIM(BOW_{CNN}, BOW)$ denotes the structural similarity image obtained by calculating the structural similarity between BOW_{CNN} and BOW. The voxel-wise structural similarity image was computed using the python package scikit-image version 0.14 using parameters to match the original SSIM definition proposed in Wang et al. (2004). The dynamic range L of the input floating-point images, needed to calculate the constants $c_1 = (0.01 L)^2$ and $c_2 = (0.03 L)^2$ in the SSIM definition, was set to twice the maximum of the reference BOW image.

In total, 5490 ROIs in 61 test subjects were analyzed. To summarize the results, the ROIs were grouped into the following composite regions: frontal cortex, temporal cortex, occipital cortex, parietal cortex, hippocampus, cingulate cortex, thalamus, basal ganglia, cerebellum, white matter, and ventricle based on their anatomical location.

2.7. Additional network robustness tests

To test the robustness of our trained CNN with respect to the noise level of the input OSEM PET images and the contrast of the input structural MR image, three additional tests were performed on two of the $[^{18}\text{F}]\text{FDG}$ data sets. First, predictions from 1 min, 3 min, and 20 min input OSEM PET images and the same T1-weighted structural MR image were compared. Second, predictions of a 20 min OSEM PET image and two types of input MRI contrasts (T1-weighted vs. FLAIR) were analyzed. Moreover, we performed predictions where we artificially degraded the image quality of the anatomical prior image by removing contrast (using a flat MR image), shifting the MR by 2.8 mm in the left-right direction, introducing a susceptibility artifact in the MR, and replacing the MR by a pseudo-CT generated from the MR (Burgos et al., 2014). Finally, we also performed a prediction for a 10 min PET/MR acquisition of the cervical spine with $[^{18}\text{F}]\text{DPA-714}$ to test the performance in an acquisition outside the brain.

3. Results

3.1. Quantitative regional analysis

Fig. 2 and supplementary Table S1 show the results for RCmean and SSIM between BOW_{CNN} and BOW evaluated in the test cases for three different PET tracers. For all three tracers and all regions except the ventricles in $[^{18}F]PE2I$, the regionally averaged RCmean ranges from 0.98 to 1.03 with a standard deviation below 0.03. For $[^{18}F]PE2I$, the ventricles show a small positive bias in the RCmean (1.08 ± 0.05). However, it should be noted that the $[^{18}F]PE2I$ uptake in the ventricles is extremely low. Only 243 out of the 5490 analyzed ROIs (4.4%) showed more than 5% bias in the RCmean. Out of those, 76 ROIs were ventricles with low tracer uptake.

The regionally averaged SSIM between BOW_{CNN} and BOW ranges from 0.958 in the thalamus to 0.974 in white matter for $[^{18}F]FDG$, from 0.963 in the basal ganglia to 0.993 in the cingulate cortex for $[^{18}F]PE2I$, and from 0.983 in the frontal cortex to 0.987 in white matter for $[^{18}F]FET$ indicating that the image quality of BOW_{CNN} is very close to BOW in all regions for all tracers. Note that no $[^{18}F]FET$ cases were used during training.

Fig. 3 shows one of the $[^{18}F]FDG$ test cases acquired on the mMR. The visual appearance of BOW_{CNN} compared to BOW is very similar with both showing better contrast between gray matter and white matter and fewer Gibbs artifacts in the cortical gray matter compared to the input OSEM image. Moreover, BOW and BOW_{CNN} show more anatomical detail compared to OSEM (e.g. differences in tracer uptake in the external capsule/claustrium between the insula and the putamen indicated by the red arrow).

In the coronal slice, the target BOW shows small fringing artifacts (indicated by the blue arrow) that are due to suboptimal alignment of the structural MRI that was used during the anatomically-guided iterative PET reconstruction using the Bowsher prior. Those artifacts are strongly reduced in BOW_{CNN} .

Supplementary Figure S3 shows an $[^{18}F]FDG$ acquisition with the SIGNA PET/MRI scanner. Again, the visual appearance and regional quantification of BOW_{CNN} is very similar to BOW, although the noise level and the noise correlations of the input OSEM image are slightly different compared to the mMR example due to lower injected activity and the availability of time-of-flight information during reconstruction. Note that 3 $[^{18}F]FDG$ data sets from the SIGNA also were used during training.

Figs. 4 and 5 show two examples of network predictions for non-FDG data sets ($[^{18}F]PE2I$ and $[^{18}F]FET$) acquired on the SIGNA. Again for both tracers, the similarity in the visual appearance between BOW_{CNN} and BOW is high. Moreover, BOW_{CNN} and BOW both show much lower noise levels compared to OSEM while preserving details and anatomical boundaries. While these results might be expected for the $[^{18}F]PE2I$ cases, the promising results for the $[^{18}F]FET$ cases are rather surprising since, first, no $[^{18}F]FET$ data sets were used during training, and second, the regional contrasts in the $[^{18}F]FET$ cases are substantially different from the $[^{18}F]FDG$ and $[^{18}F]PE2I$ cases used for training. Moreover, supplementary Fig. S1 shows the result of our trained CNN applied to a single $[^{11}C]PIB$

image (20 min acquisition, 42 min p.i., 351 MBq) where very similar promising results were obtained in a control subject without amyloid binding (only aspecific white matter binding is seen).

3.2. Additional network robustness tests

Fig. 6 shows predictions of the same [^{18}F]FDG data set as in Fig. 3 for different noise levels of the input OSEM image originating from reconstructions of 20 min, 3 min, and 1 min of emission data. In terms of regional contrast and detail, the predicted BOW_{CNN} images are very similar to the target (the reconstructed Bowsher image from 20 min of emission data). However, it can be seen that a noise cluster visible in the anterior right putamen of the 1 min and 3 min OSEM images, indicated by the red arrows, leads to a small focus with slightly increased signal in the BOW_{CNN} images predicted from the 1 min and 3 min OSEM images which is not seen in the BOW_{CNN} nor the BOW from the 20 min data. All in all, this additional test indicates that in general the output of the network is very robust against the input noise level of the OSEM images which was achieved by using data sets with different noise level during training.

Supplementary Figure S4 shows two predictions using the same [^{18}F]FDG input OSEM image with two different MRI contrasts (T1 weighted and FLAIR) and the reconstructed Bowsher images using the T1-weighted and the FLAIR MR image as structural prior image. Comparing the BOW_{CNN} images predicted from T1-weighted and FLAIR MR images, it is clear that while the overall image quality is quite similar, the gyri in the FLAIR BOW_{CNN} are slightly wider and softer. However, this is also the case for the reconstructed Bowsher image using the FLAIR MRI as structural prior image (shown in the bottom right) and can be attributed to the slightly wider and more homogeneous appearance of the gyri in the FLAIR MRI compared to the T1-weighted MRI, which is a direct result of the effects of T2-decay during the much longer readout lengths used for a FLAIR sequence which is a Fast Spin Echo based sequence. In contrast to the T1 BOW_{CNN} image, the FLAIR BOW_{CNN} image shows the locally increased uptake in the red nuclei in the midbrain much better as indicated by the red arrows. This might be because red nuclei show only very low contrast in the T1-weighted MRI while having very good contrast in the FLAIR MRI due to their high iron content.

The results shown in supplementary Figure S2 demonstrate that the output of our trained CNN is very close to the input OSEM image when no MR information (constant MR image as input) is used. In this case, the prediction actually shows increased Gibbs overshoots similar to an iterative deblurring without any regularization. This, in turn, indicates that the trained CNN did not learn specific anatomical pattern seen in the input OSEM PET images of the training cohort, but rather performs an anatomically-guided deconvolution and denoising.

Supplementary Figs. S5, S6, S7, and S8 demonstrate the amount and structure of the bias introduced by artificially degrading the image quality of the anatomical prior image in several ways is very comparable between BOW and BOW_{CNN} . Note that BOW_{CNN} slightly ameliorates high frequency artifacts introduced by misregistration. Supplementary Fig. S10 shows that also for a non-brain image, the image quality of BOW and BOW_{CNN} are very

similar. However, in this non-brain case acquired with another tracer not used during training, BOW_{CNN} is slightly smoother compared to BOW.

3.3. Prediction time

Table 3 shows the prediction time of the trained network as a function of the input size and for an Intel(R) Xeon(R) CPU E5-2699 v4 and a NVIDIA Tesla P100 SXM2 GPU. For typical brain image sizes of (220,220,220) when using a $1 \times 1 \times 1 \text{ mm}^3$ voxel grid, the prediction time is roughly 1 min on a Intel Xeon E5-2699 CPU and 1 s on a single Nvidia Tesla P100 GPU.

4. Discussion

The results obtained in this proof of concept study demonstrate that a purely image-based shift-invariant CNN can be used to approximate the results from anatomically-guided brain PET reconstruction using the asymmetric Bowsler prior. The predictions of our trained network preserve regional quantification compared to the anatomically-guided PET images obtained from iterative model-based reconstruction. Moreover, the similarity in the image quality between the predictions and the ground truth (the anatomically-guided PET reconstruction obtained with model-based iterative reconstruction) is also very high, as shown in our examples and quantified by the regional structural similarity (see Fig. 2).

Interestingly, our network also performed well on tracers not used during training, such as the seven test cases acquired with [^{18}F]FET and the single case acquired with [^{11}C]PIB (see Figs. 2 and 5, and supplementary Fig. S1). Those results and the fact that the local contrast of [^{18}F]FET and [^{11}C]PIB is different from the two tracers used during training, indicates that the network did not learn the typical local contrasts and uptake patterns seen during training, but rather learned how to perform the anatomically-constrained deconvolution and denoising of the input OSEM PET images.

As shown in Fig. 6 and supplementary Fig. S4, the image quality of the predicted images of our trained network is robust against the input MRI contrast and the noise level of the input PET images. Both of these desirable properties were achieved by using dedicated data augmentation techniques during training, namely, random contrast flipping of the training MR images and training on OSEM PET images with different count (noise) levels. Moreover, training on data from two different PET/MRI scanners resulted in a trained network that is robust against small differences in the point spread function and noise correlations of the input OSEM PET images. Note that in our CNN training, BOW images reconstructed from the highest available counts (20 min) with a fixed prior strength were used as target, even for input OSEM images reconstructed from lower count data (1 min, 3 min). In that sense, the behavior of BOW_{CNN} is different from BOW for low count data where the latter will lead to more noise in the image for a fixed prior strength. As shown in Fig. 6, this training strategy leads to a CNN whose output noise level depends less on the input noise level of the OSEM images. We believe that this training strategy is desirable for clinical applications. Moreover, we are convinced that by changing the target BOW images during training, a CNN that replicates the exact behavior of BOW in terms of bias and noise for a fixed prior strength also for low count data could be trained if needed.

Note that the differences in the point spread function and noise correlations between the OSEM images from the SIGNA and mMR are due to slightly different crystal sizes, the availability of ca. 400 ps time of flight resolution in the SIGNA and the number of iterations and subsets used in the OSEM reconstructions. Overall these results suggest that our trained model should also perform well on tracers that we have not tested in this proof of concept work. More extensive validation of this hypothesis will be part of future work.

Using a purely convolutional shift-invariant network allowed us to train the network on small 3D patches. This in turn meant that many small training patches carrying different information could be obtained from a single PET/MRI acquisition which enabled us to train the network on a relatively small number of PET/MRI acquisitions (26 in total). This property is important since to the best of our knowledge large pooled PET/MRI databases containing PET emission raw data and structural MRI data, that can be used to generate large sets of training pairs, do not exist so far. Undoubtedly, we expect that the performance of the network can be improved even further by using more training data sets. Note that with our implementation, training of our network took ca. 8.5 h such that retraining on new data can be performed overnight.

We believe that the very fast prediction time (ca. 1 s on a recent GPU) is the biggest advantage of using a convolutional neural network to predict anatomically-guided PET reconstructions in image space. This means that, first, there is no clinically meaningful time delay to generate an anatomically-guided PET reconstruction next to the standard OSEM PET images. Second, those ultra-fast prediction times allow the generation of anatomically-guided PET reconstructions with different level of regularization (from different trained networks) in almost real-time. Consequently, it would be possible to quickly switch between PET images with different levels of MRI-based regularization which will help clinical readers to understand the influence of MRI-based regularization and, moreover, help to tune the level of regularization for different local clinical applications. Because our network can be applied retrospectively on reconstructed DICOM images, this will facilitate the execution of bigger multi-center studies to assess the value of anatomically-guided PET images for different clinical applications.

5. Limitations

Our proof of concept study has a few limitations. First of all, the similarity in image quality between the anatomically-guided PET images obtained from the network prediction and model-based iterative reconstruction was assessed by two mathematical quantities and not by an observer study using a well defined clinical application. However, given the very small differences documented by these measures and the improved sensitivity of the anatomically-guided PET reconstruction for the detection of pathology (Shepherd et al., 2019) we foresee that the clinical benefits of our proposed CNN approach will be readily apparent. We plan to present the results of an observer study evaluating the performance of anatomically-guided PET images predicted from our network in the context of brain [18F]FDG in epilepsy in upcoming work. Second, in this work, we only evaluated a network trained on anatomically-guided PET reconstructions using one specific type of segmentation-free MRI-based priors, namely, the asymmetric Bowsher prior. This prior was chosen based on a previous study

(Schramm et al., 2017), where we could show that it has slightly superior performance compared to the parallel level sets prior (Ehrhardt et al., 2016). However, even though we did not investigate additional priors in our study, we expect that a network to predict PET images reconstructed with a different prior (such as e.g. parallel level sets) can attain similar performance. Last but not least, in this work we only trained and thoroughly tested our network on brain images. Extensive validation of applications of the network to regions outside the brain / head will be part of future work. Until then, we would recommend to use the CNN of this work only trained on brain images only for prediction of brain acquisitions. Third, the performance of the trained network on different tracers is of course influenced by the number and balance between tracers or PET contrasts used during training. In our case, the training data set was dominated by FDG images. To achieve even better performance for an application using a specific single tracer only, retraining the network with different tracers might be desirable.

6. Conclusion

We demonstrated that anatomically-guided model-based iterative PET reconstruction can be well-approximated by a purely shift-invariant convolutional neural network in image space allowing the generation of anatomically-guided PET images in almost real-time. Using dedicated data augmentation techniques during training resulted in a CNN that is robust against the used PET tracer, the noise level of the input PET images and the input MRI contrast.

7. Data and code availability

The trained network, the implementation of the network training and examples on how to do predictions (from clinical DICOM and nifti images) will be provided online at <https://gschramm.github.io/pyapetnet/> and can be used for non-commercial research purposes upon acceptance of this article for publication. The training, validation and test data sets cannot be made available due to privacy protection restrictions of our institutions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

GS is funded by the NIH project 1P41EB017183-01A1. AR and JN are funded by the Research Foundation Flanders (FWO) projects 12T7118N and G.0275.14N.

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X, 2015 Tensorflow: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org <http://tensorflow.org/>.

- Baete K, Nuyts J, Paesschen WV, Suetens P, Dupont P, 2004 Anatomical-based FDG-PET reconstruction for the detection of hypo-metabolic regions in epilepsy. *IEEE Trans. Med. Imaging* 23 (4), 510–519. doi:10.1109/TMI.2004.825623. [PubMed: 15084076]
- Bland J, Mehranian A, Belzunce MA, Ellis S, McGinnity CJ, Hammers A, Reader AJ, 2017 MR-guided kernel EM reconstruction for reduced dose PET imaging. *IEEE Trans. Radiat. Plasma Med. Sci.* 2 (3). doi:10.1109/trpms.2017.2771490. 1–1. [PubMed: 29930991]
- Bowsher JE, Yuan H, Hedlung LW, Turkington TG, Akabani G, Badea A, Kurylo WC, Wheeler CT, Cofer GP, Dewhirst MW, Johnson A, 2004 Utilizing MRI information to estimate f18-FDG distributions in rat flank tumors. In: *Proceedings of the Nuclear Science Symposium Conference Record IEEE* 4 (C), pp. 2488–2492. doi:10.1109/NSSMIC.2004.1462760. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1462760.
- Burgos N, Cardoso MJ, Thielemans K, Modat M, Pedemonte S, Dickson J, Barnes A, Ahmed R, Mahoney CJ, Schott JM, Duncan JS, Atkinson D, Arridge SR, Hutton BF, Ourselin S, 2014 Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies. *IEEE Trans. Med. Imaging* 33 (12), 2332–2341. doi:10.1109/TMI.2014.2340135. [PubMed: 25055381]
- Chen KT, Gong E, Bezerra F, Macruz DC, Xu J, 2019 Ultra - low-dose 18 fflorbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. *Radiology* 290, 649–656. doi:10.1148/radiol.2018180940. [PubMed: 30526350]
- Chollet F, et al., 2015 Keras. <https://keras.io>.
- Comtat C, Kinahan PE, Fessler JA, Beyer T, Townsend DW, Defrise M, Michel C, 2002 Clinically feasible reconstruction of 3d whole-body PET/CT data using blurred anatomical labels. *Phys. Med. Biol.* 47 (1), 1–20. doi:10.1088/0031-9155/47/1/301. [PubMed: 11814220]
- Cui J, Liu X, Wang Y, Liu H, 2017 Deep reconstruction model for dynamic PET images. *PLoS ONE* 12 (9), 1–21. doi:10.1371/journal.pone.0184667.
- De Man B, Basu S, 2004 Distance-driven projection and backprojection in three dimensions. *Phys. Med. Biol.* 49 (11), 2463–2475. doi:10.1109/NSSMIC.2002.1239600. [PubMed: 15248590]
- Delso G, Furst S, Jakoby B, Ladebeck R, Ganter C, Nekolla SG, Schwaiger M, Ziegler SI, 2011 Performance measurements of the siemens mMR integrated whole-body PET/MR scanner. *J. Nucl. Med.* 52 (12), 1914–1922. doi:10.2967/jnumed.111.092726. [PubMed: 22080447]
- Dong C, Loy CC, He K, Tang X, 2016 Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intel.* 38 (2), 295–307. doi:10.1109/TPAMI.2015.2439281. <http://www.ncbi.nlm.nih.gov/pubmed/26761735>.
- Ehrhardt M, Markiewicz P, Liljeroth M, Barnes A, Kolehmainen V, Duncan J, Pizarro L, Atkinson D, Hutton BF, Ourselin S, Thielemans K, Arridge SR, 2016 PET reconstruction with an anatomical MRI prior using parallel level sets. *IEEE Trans. Med. Imaging* 35 (9), 2189–2199. doi:10.1109/TMI.2016.2549601. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7452643>. [PubMed: 27101601]
- Erlandsson K, Buvat I, Pretorius PH, Thomas BA, Hutton BF, 2012 A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys. Med. Biol.* 57 (21). doi:10.1088/0031-9155/57/21/R119.
- Fazio F, Perani D, 2000 Importance of partial-volume correction in brain PET studies. *J. Nucl. Med.: Off. Publ. Soc. Nucl. Med.* 41 (11), 1849–1850. <http://www.ncbi.nlm.nih.gov/pubmed/11079493&percent;5Cn11079493>.
- Fischl B, Kouwe AVD, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM, 2004 Automatically parcellating the human cerebral cortex. *Cerebral Cortex* 14 (1), 11–22. doi:10.1093/cercor/bhg087. [PubMed: 14654453]
- Gong K, Catana C, Qi J, Li Q, 2019 PET image reconstruction using deep image prior. *IEEE Trans. Med. Imaging* 38 (7), 1655–1665. doi:10.1109/TMI.2018.2888491. [PubMed: 30575530]
- Gong K, Guan J, Kim K, Zhang X, Yang J, Seo Y, Fakhri GE, Qi J, Li Q, 2019 Iterative PET image reconstruction using convolutional neural network representation. *IEEE Trans. Med. Imaging* 38 (3), 675–685. doi:10.1109/TMI.2018.2869871. [PubMed: 30222554]
- Grant AM, Deller TW, Khalighi MM, Maramraju SH, Delso G, Levin CS, 2016 NEMA NU 2–2012 performance studies for the siPM-based tof-PET component of the GE SIGNA PET/MR system.

- Med. Phys. 43, 2334–2343. doi:10.1118/1.4945416. <http://scitation.aip.org/content/aapm/journal/medphys/43/5/10.1118/1.4945416>. [PubMed: 27147345]
- Hägström I, Schmidlein CR, Campanella G, Fuchs TJ, 2019 Deep PET : a deep encoder - decoder network for directly solving the PET image reconstruction inverse problem. *Med. Image Anal.* 54, 253–262. doi:10.1016/j.media.2019.03.013. [PubMed: 30954852]
- Hong X, Zan Y, Weng F, Tao W, Peng Q, Huang Q, 2018 Enhancing the image quality via transferred deep residual learning of coarse PET sinograms. *IEEE Trans. Med. Imaging* 37 (10), 2322–2332. doi:10.1109/TMI.2018.2830381. [PubMed: 29993685]
- Hutton BF, Thomas BA, Erlandsson K, Bousse A, Reilhac-Laborde A, Kazantsev D, Pedemonte S, Vunckx K, Arridge SR, Ourselin S, 2013 What approach to brain partial volume correction is best for PET/MRI? In: *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Vol. 702, pp. 29–33. doi:10.1016/j.nima.2012.07.059. 10.1016/j.nima.2012.07.059.
- Kim K, Wu D, Gong K, Dutta J, Kim JH, Son YD, Kim HK, Fakhri GE, Li Q, 2018 Penalized PET reconstruction using deep learning prior and local linear fitting. *IEEE Trans. Med. Imaging* 37 (6), 1478–1487. doi:10.1109/TMI.2018.2832613. [PubMed: 29870375]
- Knoll F, Holler M, Koesters T, Otazo R, Bredies K, Sodickson DK, 2016 Joint MR-PET reconstruction using a multi-channel image regularizer. *IEEE Trans. Med. Imaging* 36 (1), 1–16. doi:10.1109/TMI.2016.2564989.
- Lipinski B, Herzog H, Kops ER, Oberschelp W, Muller-Gartner HW, 1997 Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. *IEEE Trans. Med. Imaging* 16 (2), 129–136. doi:10.1109/42.563658. <http://www.ncbi.nlm.nih.gov/pubmed/9101322>. [PubMed: 9101322]
- Liu P, Zhang H, Zhang K, Lin L, Zuo W, 2018 Multi-level wavelet-CNN for image restoration. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 886–895. doi:10.1109/CVPRW.2018.00121.
- Loeb R, Navab N, Ziegler S, 2015 Direct parametric reconstruction using anatomical regularization for simultaneous PET/MRI data. *IEEE Trans. Med. Imaging* 34 (11), 2233–2247. doi:10.1109/TMI.2015.2427777. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7097704>. [PubMed: 25935030]
- Mehranian A, Belzunce MA, McGinnity CJ, Bustin A, Prieto C, Hammers A, Reader AJ, 2018 Multi-modal synergistic PET and MR reconstruction using mutually weighted quadratic priors. *Magn. Resonance Med.* (May) 1–15. doi:10.1002/mrm.27521.
- Mueller-Gaertner HW, Links JM, Prince JL, Brakeman PR, Bryan RN, Mcveigh E, Leal JP, Davatzikos C, Frost JJ, 1992 Measurement of radiotracer concentration in brain gray-matter using PET (GM-PET) - MRI-based correction for partial volume effect. *J. Cerebral Blood Flow Metabol.* 19 (8), 645.
- Nuyts J, 2007 The use of mutual information and joint entropy for anatomical priors in emission tomography. *IEEE Nucl. Sci. Symp. Conf. Record* 6 (2), 4149–4154. doi:10.1109/NSSMIC.2007.4437034.
- Nuyts J, Bequé D, Dupont P, Mortelmans L, 2002 A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Trans. Nucl. Sci.* 49 (1), 56–60. doi:10.1109/TNS.2002.998681.
- Rigie DS, Schramm G, Vahle T, Shepard T, Nuyts J, Boada FE, 2018 Approximating MRI-based anatomically guided PET reconstruction with a convolutional neural network. In: *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) IEEE*, pp. 1–3.
- Schramm G, Holler M, Rezaei A, Vunckx K, Knoll F, Bredies K, Boada F, Nuyts J, 2017 Evaluation of parallel level sets and bowsher’s method as segmentation-free anatomical priors for time-of-flight PET reconstruction. *IEEE Trans. Med. Imaging* 37 (2), 590–603. doi:10.1109/TMI.2017.2767940.
- Shepherd T, Schramm G, Vahle T, Rigie D, Friedman K, Zan E, Ghesani M, Nuyts J, Boada F, 2019 Impact of MR-guided PET reconstruction on seizure foci localization with FDG PET. *J. Nucl. Med.* 392. 60 (supplement 1) http://jnm.snmjournals.org/content/60/supplement_1/392.abstract.

- Ulyanov D, Vedaldi A, Lempitsky V, 2018 Deep image prior. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition IEEE, pp. 9446–9454. doi:10.1109/CVPR.2018.00984. ArXiv:1711.10925.
- Vunckx K, Atre A, Baete K, Reilhac A, Deroose CM, Laere KV, Nuyts J, 2012 Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging. *IEEE Transactions on Medical Imaging* 31 (3), 599–612. doi:10.1109/TMI.2011.2173766. [PubMed: 22049363]
- Vunckx K, Dupont P, Goffin K, Paesschen WV, Laere KV, Nuyts J, 2014 Voxel-based comparison of state-of-the-art reconstruction algorithms for 18f-FDG PET brain imaging using simulated and clinical data. *NeuroImage* 102 (P2), 875–884. doi:10.1016/j.neuroimage.2014.06.068. [PubMed: 25008958]
- Vunckx K, Nuyts J, 2010 Heuristic modification of an anatomical Markov prior improves its performance. In: Nuclear Science Symposium Conference Record, 2010 IEEE, pp. 1–5.
- Wang Y, Yu B, Wang L, Zu C, Lalush DS, Lin W, Wu X, Zhou J, Shen D, Zhou L, 2018 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *NeuroImage* 174 (2017), 550–562. doi:10.1016/j.neuroimage.2018.03.045. [PubMed: 29571715]
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP, 2004 Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612. [PubMed: 15376593]
- Xiang L, Qiao Y, Nie D, An L, Lin W, Wang Q, Shen D, 2017 Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. *Neurocomputing* 267, 406–416. doi:10.1016/j.neucom.2017.06.048. [PubMed: 29217875]
- Yang B, Ying L, Tang J, 2018 Artificial neural network enhanced Bayesian PET image reconstruction. *IEEE Trans. Med. Imaging* 37 (6), 1297–1309. doi:10.1109/TMI.2018.2803681. [PubMed: 29870360]
- Yang J, Huang S, Mega M, Lin K, Toga A, Small G, Phelps M, 1996 Investigation of partial volume correction methods for brain FDG PET studies. *IEEE Trans. Nucl. Sci.* 43 (6), 3322–3327. doi:10.1109/23.552745.
- Zhang K, Zuo W, Chen Y, Meng D, Zhang L, 2018 Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* 26 (7), 3142–3155. doi:10.1109/TIP.2017.2662206. <http://arxiv.org/abs/1802.08808>.

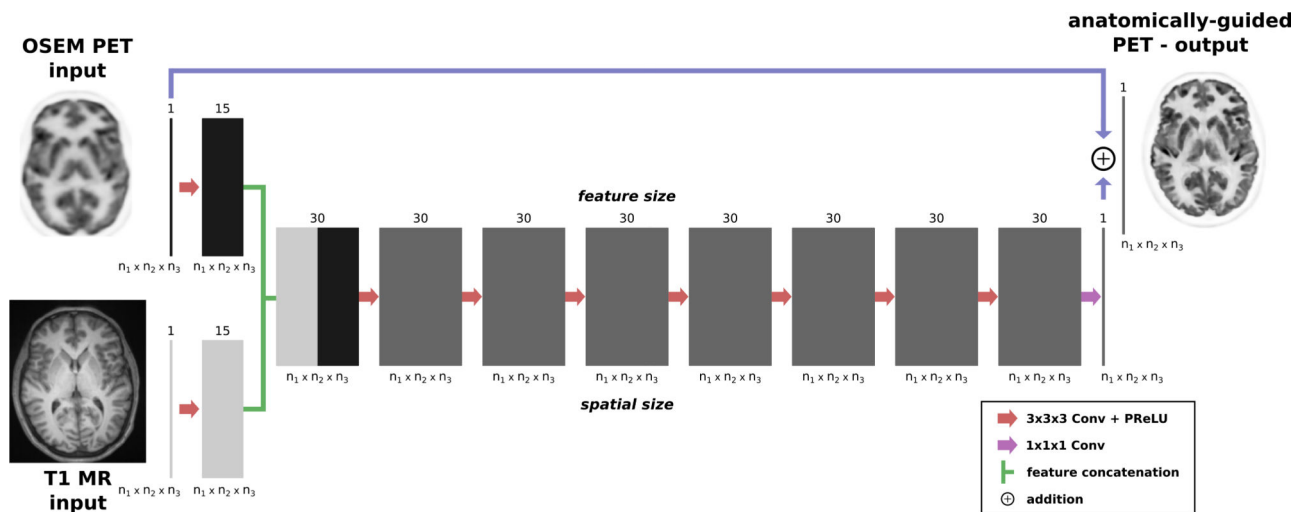


Fig. 1. Architecture of our convolutional neural network to predict a 3D anatomically-guided PET reconstructions from an input 3D OSEM PET image and a 3D structural MR image. See text for details.

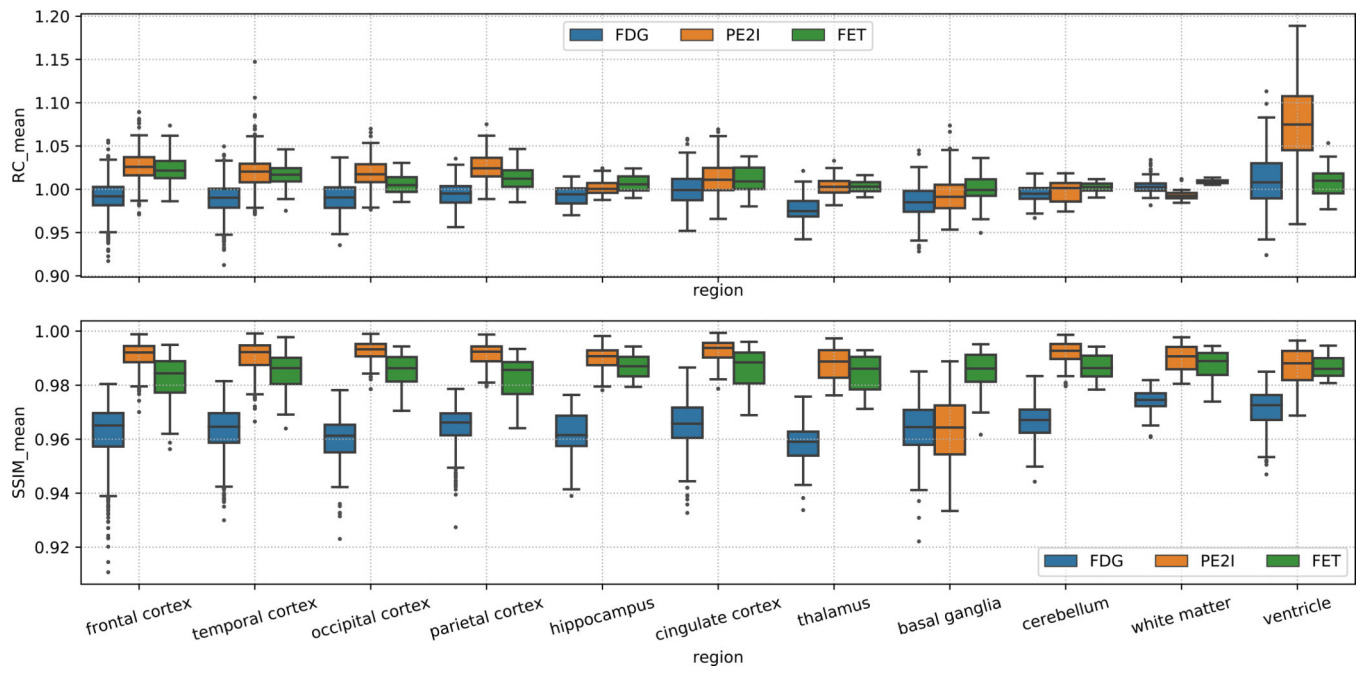


Fig. 2. Boxplots of regional values for RC_{mean} (top) and $SSIM_{\text{mean}}$ (bottom) between the BOW_{CNN} and BOW in the $[^{18}\text{F}]FDG$ (blue), $[^{18}\text{F}]PE2I$ (orange), and $[^{18}\text{F}]FET$ (green) test cases.

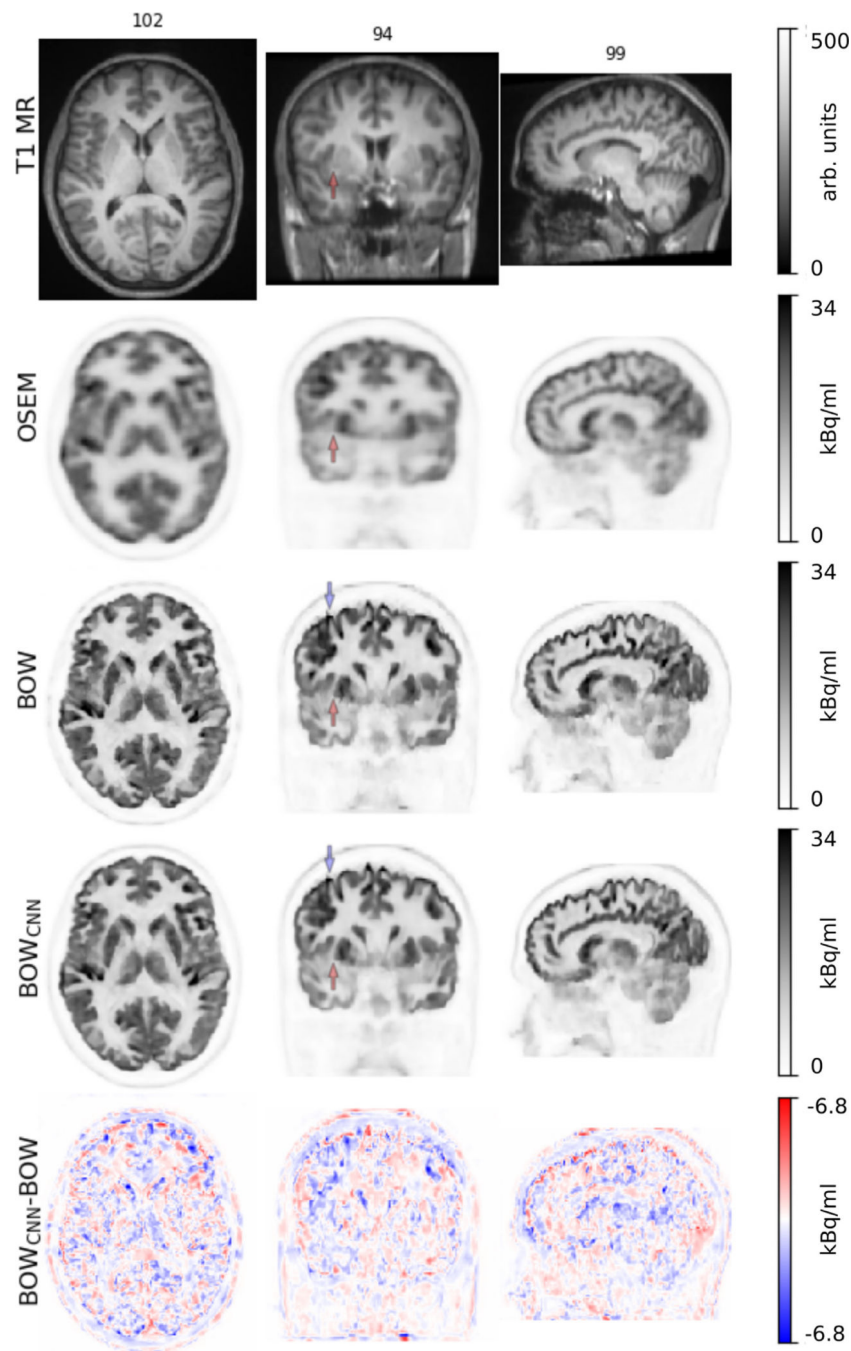


Fig. 3. Example [^{18}F]FDG test case acquired on the mMR. (top row) structural T1-weighted MRI used as prior image in the iterative anatomically-guided PET reconstruction using the Bowsher prior. (2nd row) Standard OSEM PET reconstruction obtained from 20 min emission data. (3rd row) reference iterative anatomically-guided PET reconstruction using the Bowsher prior (BOW). (4th row) prediction of our trained convolutional neural network (BOW_{CNN}) using the OSEM PET image and the structural MRI as input. (5th row) absolute difference between BOW_{CNN} and BOW. The red arrow indicates the location of the right

claustrum between the insula and putamen where BOW and BOW_{CNN} show more anatomical detail compared to OSEM. The blue arrow shows a region of fringing artifacts in BOW that less apparent in BOW_{CNN}.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

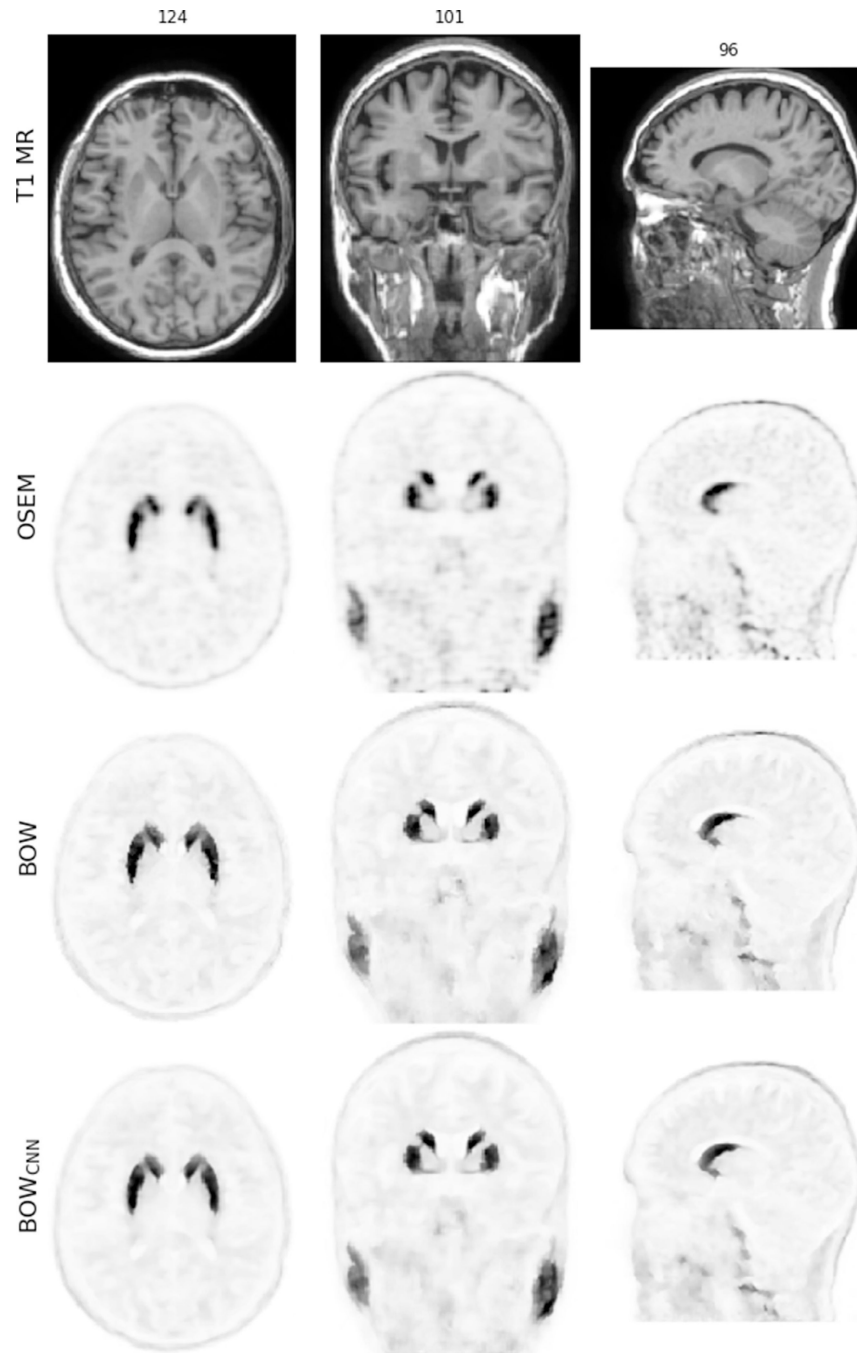


Fig. 4. Same as Fig. 3 for a $[^{18}\text{F}]\text{PE2I}$ test case acquired on the SIGNA.

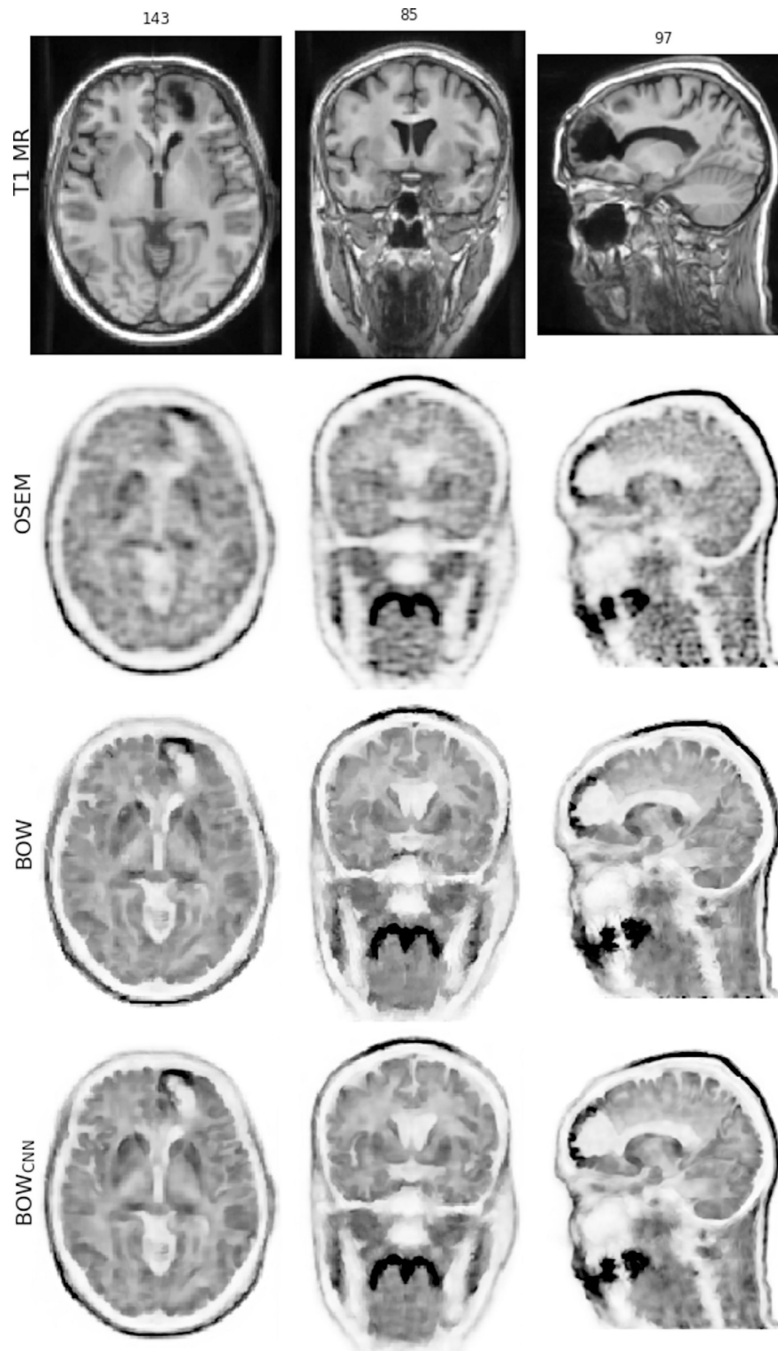


Fig. 5. Same as Fig. 3 for a $[^{18}\text{F}]\text{FET}$ test case acquired on the SIGNA. In this case the acquisition time was 25 min.

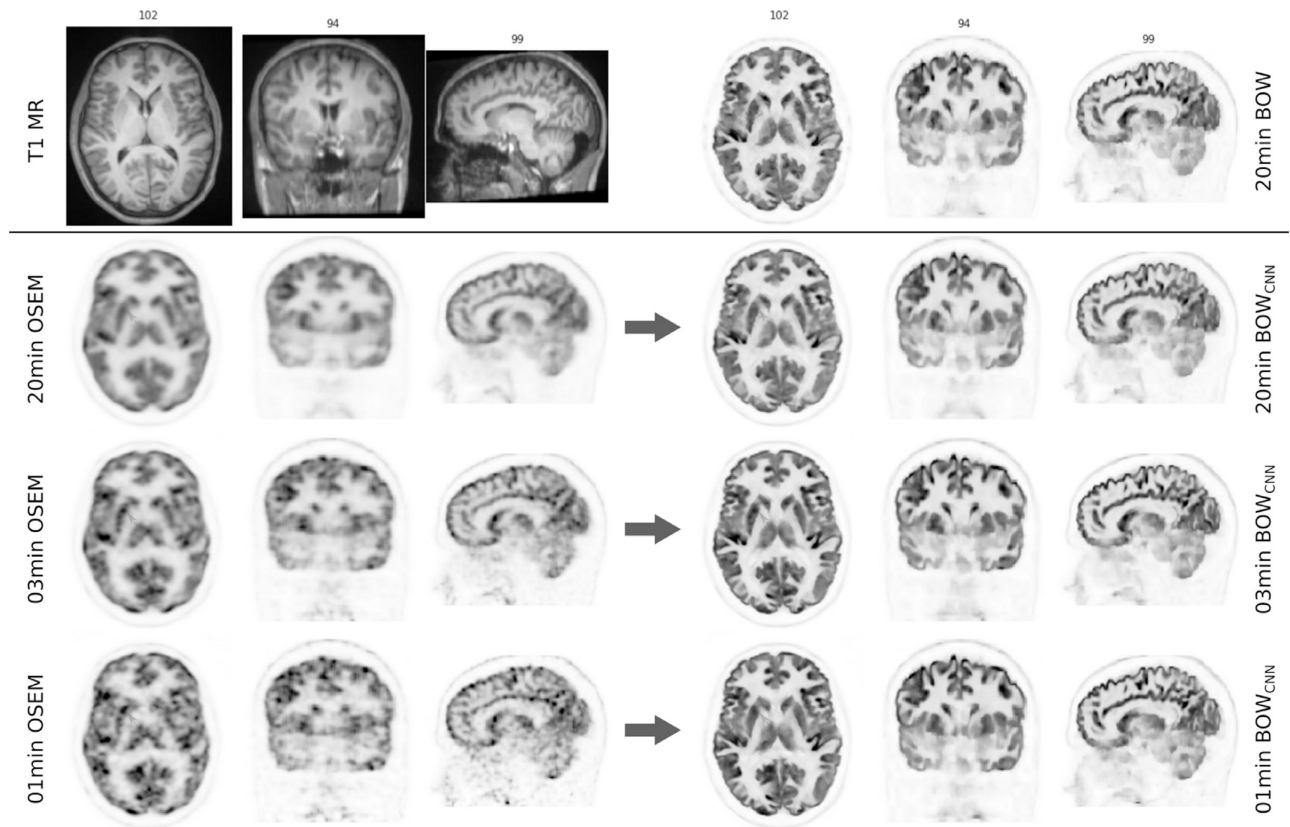


Fig. 6.

Impact of noise level in the input OSEM PET image on the image quality of the predicted anatomically-guided PET image (BOW_{CNN}). The case shown here is the same as in Fig. 3. (top row left) structural T1-weighted MRI used as prior image in the iterative anatomically-guided PET reconstruction using the Bowsher prior. (top row right) reference iterative anatomically-guided PET reconstruction using the Bowsher prior (BOW). (2nd till 4th row left) OSEM PET reconstruction obtained from 20 min, 3 min, and 1 min of emission data. (2nd till 4th row left) corresponding predictions of our trained convolutional neural network (BOW_{CNN}) using the respective OSEM PET image and the structural MRI as input. Note that although the noise level of the input OSEM images varies a lot, the noise level and the level of detail in the BOW_{CNN} images is remarkably constant and comparable to the BOW image of the full 20 min emission data. The red arrows indicate a noise cluster in the 1 min and 3 min OSEM images that leads to a small focus with slightly increased signal in the BOW_{CNN} images predicted from the those OSEM images which is not seen in the BOW_{CNN} nor the BOW from the 20 min data.

Table 1

Summary of subjects and acquisitions used for training, validation, and testing of our network.

Cohort	Tracer	Scanner	n	Acquisition time	Frame center
Training	^{18}F FDG	mMR	13	1 min, 3 min, 20 min	50 min p.i.
		SIGNA	3	20 min	50 min p.i.
	^{18}F PE2I	SIGNA	10	20 min	50 min p.i.
Validation	^{18}F FDG	mMR	2	1 min, 3 min, 20 min	50 min p.i.
		SIGNA	1	20 min	50 min p.i.
	^{18}F PE2I	SIGNA	3	20 min	50 min p.i.
Testing	^{18}F FDG	mMR	36	20 min	50 min p.i.
	^{18}F PE2I	SIGNA	18	20 min	50 min p.i.
	^{18}F FET	SIGNA	7	25 min	72.5 min p.i.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Summary of parameters used in the OSEM and anatomically-guided Bowsher reconstructions for the mMR and SIGNA cases.

Scanner	Reconstruction	Iterations	Subsets	Voxelsize	TOF resolution	β	Resolution modeling
mMR	OSEM	3	21	$1.04 \times 1.04 \times 2.03 \text{ mm}^3$	-	-	4.5 mm Gaussian convolution in sinogram space (radial and transaxial)
	Bowsher	10	42	$1.04 \times 1.04 \times 2.03 \text{ mm}^3$	-	10	4.5 mm Gaussian convolution in sinogram space (radial and transaxial)
SIGNA	OSEM	3	28	$1.39 \times 1.39 \times 1.39 \text{ mm}^3$	450 ps	-	4.5 mm Gaussian convolution in image space
	Bowsher	10	28	$1.39 \times 1.39 \times 1.39 \text{ mm}^3$	450 ps	10	4.5 mm Gaussian convolution in image space

Table 3

Prediction time of the trained network as a function of input size and for an Intel(R) Xeon(R) CPU E5-2699 v4 and a NVIDIA Tesla P100 SXM2 GPU. Mean and standard deviation over seven predictions are given.

Input size	Prediction time (ms)	
	Intel(R) Xeon(R) CPU E5-2699 v4	NVIDIA GPU Tesla P100 SXM2
(10,10,10,2)	14 ± 1	4 ± 0.5
(50,50,50,2)	826 ± 38	15 ± 0.1
(100,100,100,2)	5170 ± 118	100 ± 0.7
(150,150,150,2)	16,900 ± 455	360 ± 3.7
(200,200,200,2)	41,100 ± 545	851 ± 5
(250,250,250,2)	77,000 ± 1680	1720 ± 10

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript