Supplemental material: Dynamic PET Image Reconstruction Using Nonnegative Matrix Factorization Incorporated With Deep Image Prior

Tatsuya Yokota*,⋆, Kazuya Kawai*, Muneyuki Sakata†, Yuichi Kimura‡, and Hidekata Hontani*
* Nagoya Institute of Technology, Nagoya, Japan, {t.yokota, hontani}@nitech.ac.jp
⋆ RIKEN Center for Advanced Intelligence Project, Tokyo, Japan
† Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan
‡ Kindai University, Wakayama, Japan

Details of the proposed model

In this supplemental document, we explain more details of the proposed model for generating spatial factors. As mentioned in the main manuscript, each spatial factor is generated by single U-Net. Since each U-Net outputs a 2D/3D image from the fixed input noise, we vectorize each output image and concatenate them to make the factor matrix $\boldsymbol{A}$ (see Figure 1). The network parameters of all U-Nets are optimized for minimizing KL divergence between observed dynamic sinogram and reconstructed dynamic sinogram. The errors are propagated from sinograms to PET images, spatial factors, and individual U-Net's parameters.

Figure 2 shows the employed architecture of U-Net for our experiments, which is almost same architecture used in [1]. The employed architecture, called as U-Net, consists of convolution, downsampling, and upsampling layers with skip-connections. The "conv(n,n)" in Figure 2 stands for the convolution layer with the kernel size of (n,n). Generally, each convolution layer is combined with batch-normalization (BN) and leaky Relu activation excluding the last convolution layer. Reflection padding is used before some convolution layer with a larger kernel to keep the same resolution.

We regard a group of convolutional layers with downsampling as the "encoder network", and a group of convolutional layers with upsampling as the "decoder network". In the encoder network, we simply used strides (2,2) for downsampling operation with a convolution layer. In the decoder network, we practically used bi-linear interpolation for upsampling operation. Skip-connection is linking the encoder and decoder networks of the same resolution with a convolution layer. In our PET reconstruction model, we restrict the range of each spatial factor as [0,1] by using the sigmoid layer before the output of U-Net.

Experimental settings

For all experiments, the reconstructed spatial factor (i.e., output of each U-Net) was a 2D-image with size of (128,128). We set the code depth as $C = 32$, and the size of input noise was (128,128,32). Input noise u was generated by uniform

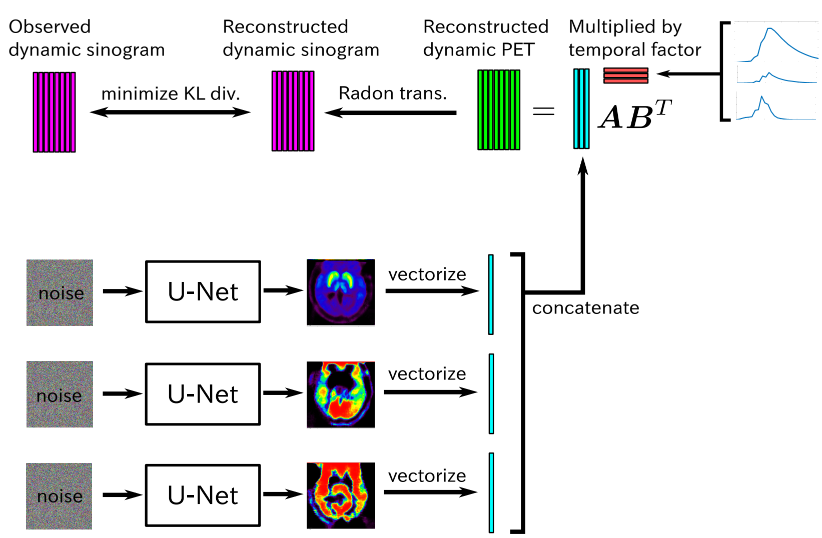


Fig. 1: Proposed reconstruction model.

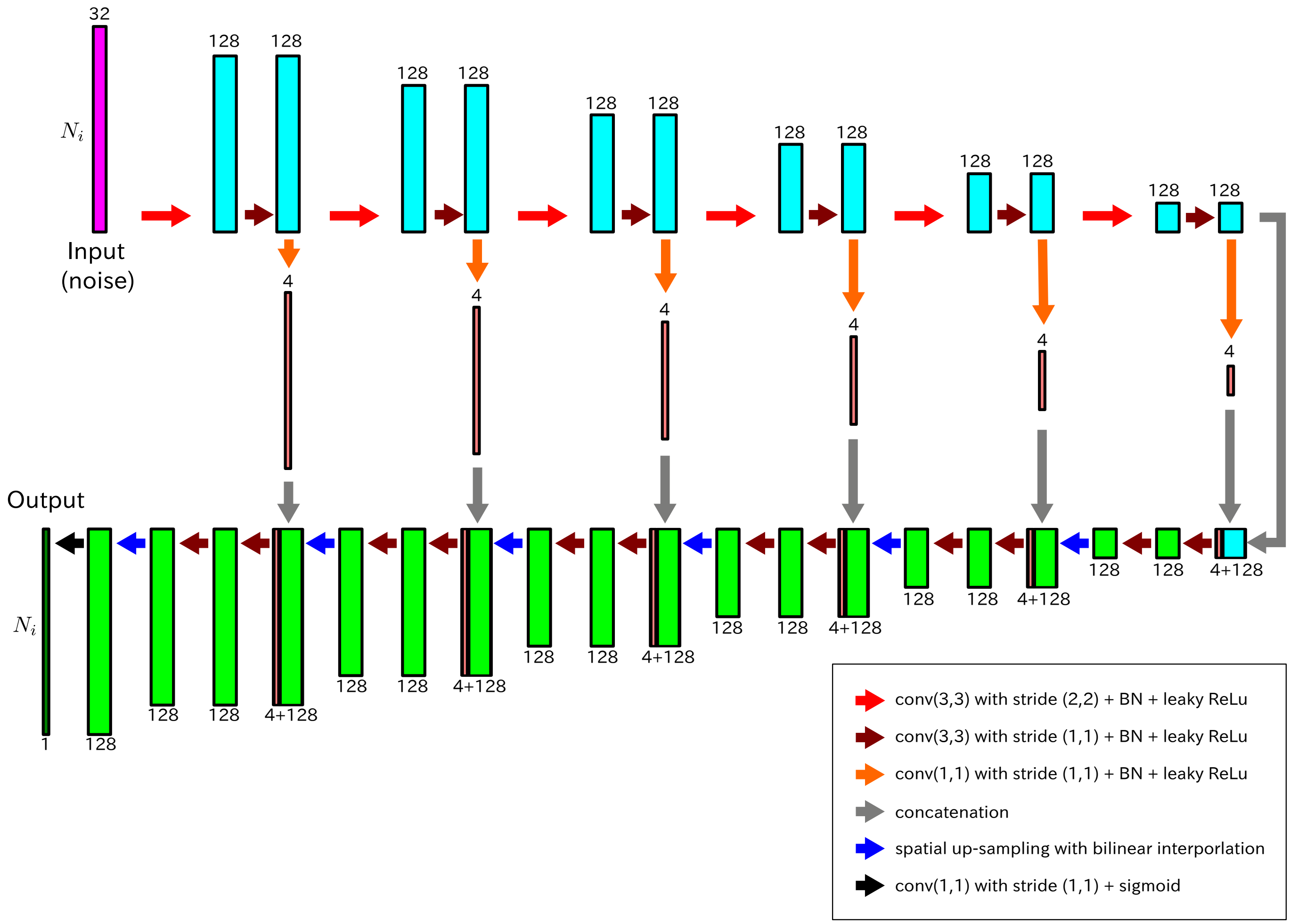


Fig. 2: Architecture of each U-Net.

distribution of range $[0, 0.1]$ which is fixed throughout the whole process, and the same noise u was inputted into the all U-Nets. Hyperparameter for the l_p-norm, we set usually $p = 0.5$, and $\alpha = 0.01$. We selectively used the value of β from $\{0.001, 0.01, 0.1, 1.0, 10.0\}$.

Preliminary experiments of single PET image reconstruction by using deep image prior (Fig.2 in the main manuscript)

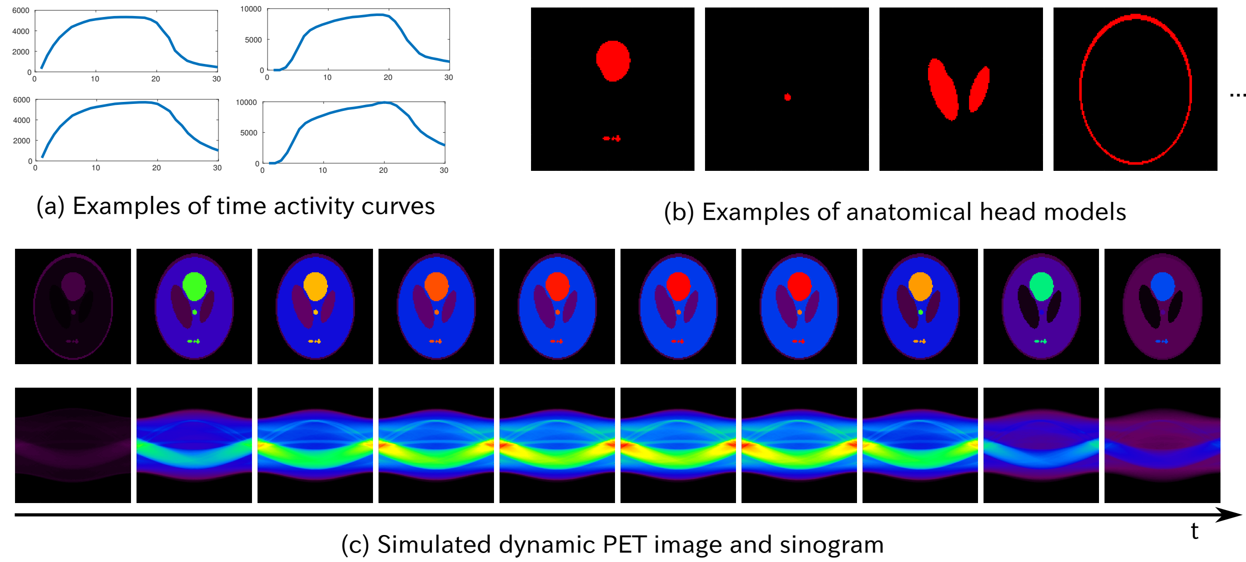
In this experiment, we used a standard Shepp-Logan phantom, which is generated from function of "phantom(128)" in `MATLAB`. The range of this phantom image was in [0,1], then it is consistent with the sigmoid output of U-Net. The sinogram was generated by the multiplication of Radon transform matrix and vectorized phantom image $\boldsymbol{Rz}$, and any noise did not be added.

For reconstruction of the Shepp-Logan phantom from sinogram, we solve the optimization problem (8) in the main manuscript with $R \in \{1, 2, ..., 5\}$. We used `TensorFlow` for implementing the proposed method. The learning rate was initially 0.01 and decayed by multiplying 0.98 every 100 steps. We iterate to update the network parameter Θ for 8000 steps via AdamOptimizer. The supplemental material includes the GIF animation of its optimization behavior.

In case of $R = 1$, we tried the various different settings of network architectures such as number of filters, with/without skip-connections, kernel sizes, however we could not find any significant difference from the default architecture. Compared with other modifications, the larger R significantly increased the degree of freedom of image representation by U-Net in our experiments.

Experiments of simulated dynamic PET image reconstruction (Fig.3-6 in the main manuscript)

In this experiment, we synthetically generated a dynamic PET image based on the compartment model and based on phantom-like anatomical head model (see Fig. 3). The different levels of noisy sinograms were generated by the Poisson resampling with scaling/unscaling. For the proposed method, we set $\mu_{\boldsymbol{B}} = 0.01$, $i_{\boldsymbol{B}} = 100$, and initial learning rate $\mu_\Theta = 0.01$. We decayed its learning rate $\mu_\Theta \leftarrow 0.98\mu_\Theta$ every 100 steps. We set $I_{\max} = 20000$ for monitoring the optimization behavior, and $I_{\max} = 5000$ for the other experiments.



(a) Examples of time activity curves

(b) Examples of anatomical head models

(c) Simulated dynamic PET image and sinogram

Fig. 3: Synthetic simulation of dynamic PET image and sinogram.

Further visualizations of experimental results

Figure 4–14 show the whole time visualizations of reconstructed dynamic PET images by using the all baseline methods.

References

[1] D. Ulyanov, A. Vedaldi, and V. Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, June 2018.

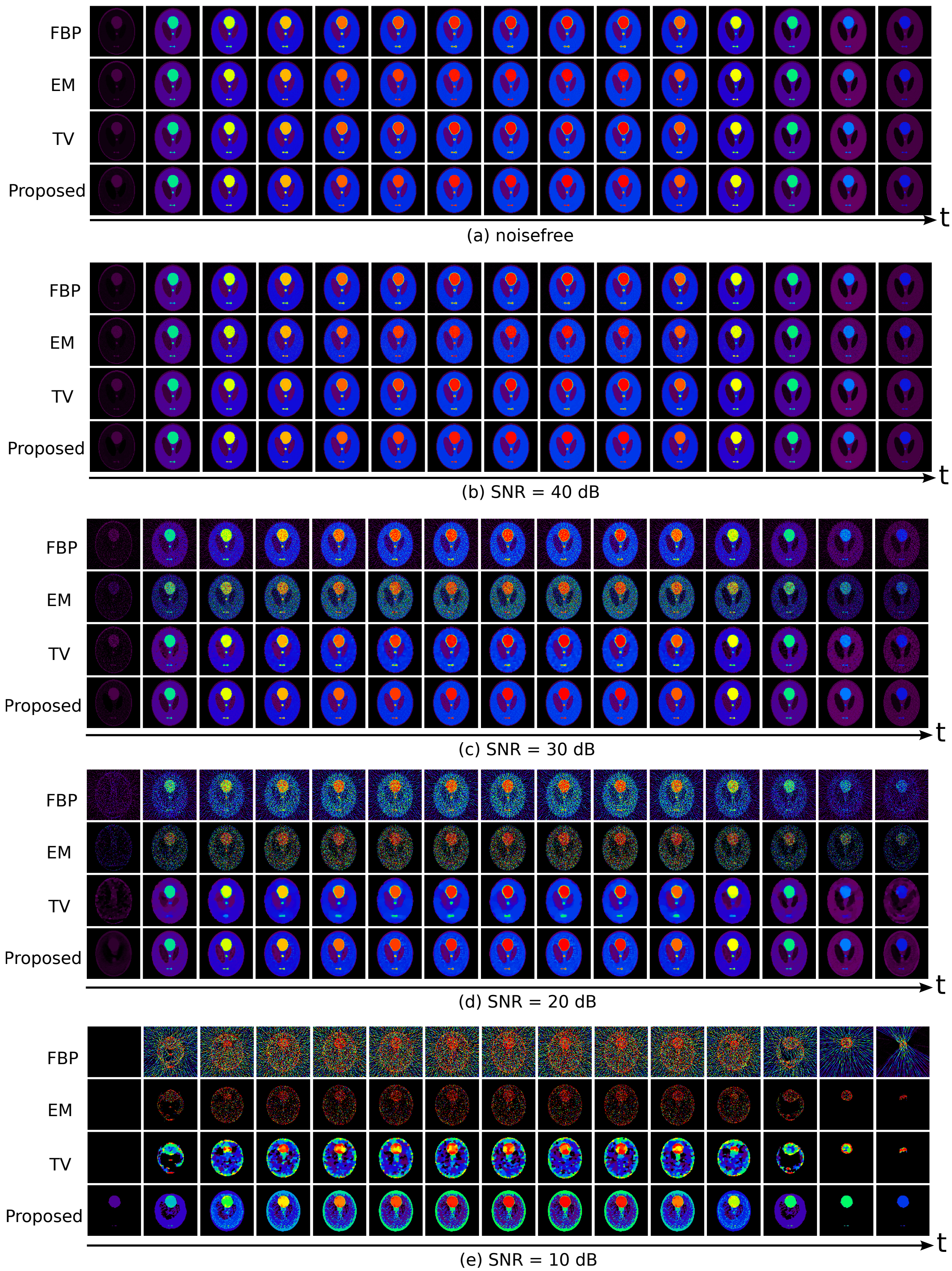


Fig. 4: Results of dynamic PET image reconstruction of phantom data.

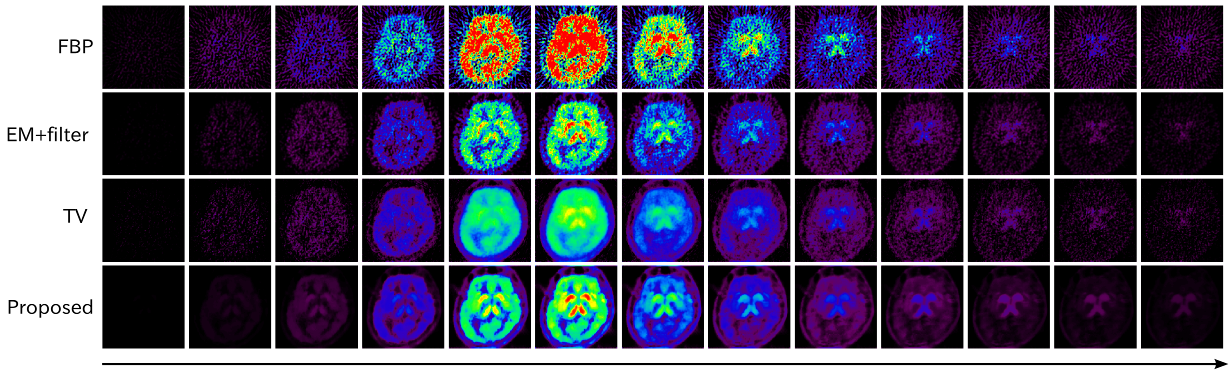


Fig. 5: Results of dynamic PET image reconstruction of clinical data: Subject 1.

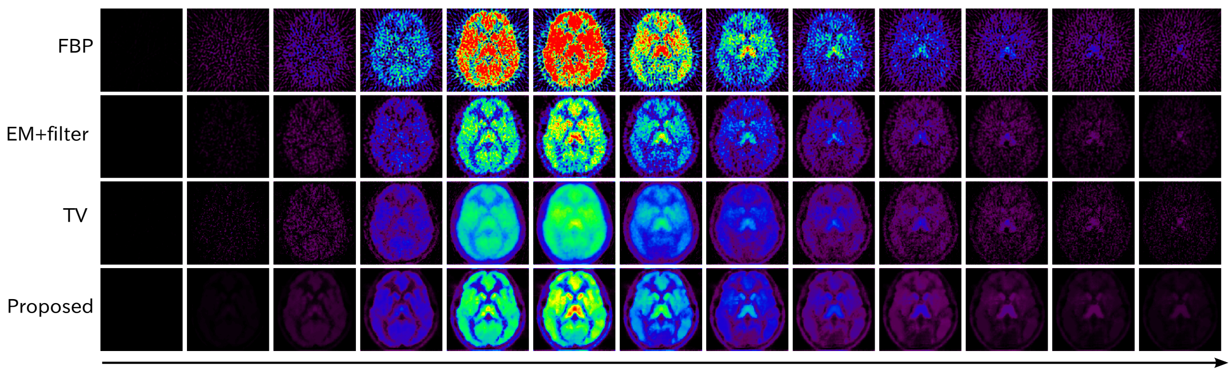


Fig. 6: Results of dynamic PET image reconstruction of clinical data: Subject 2.

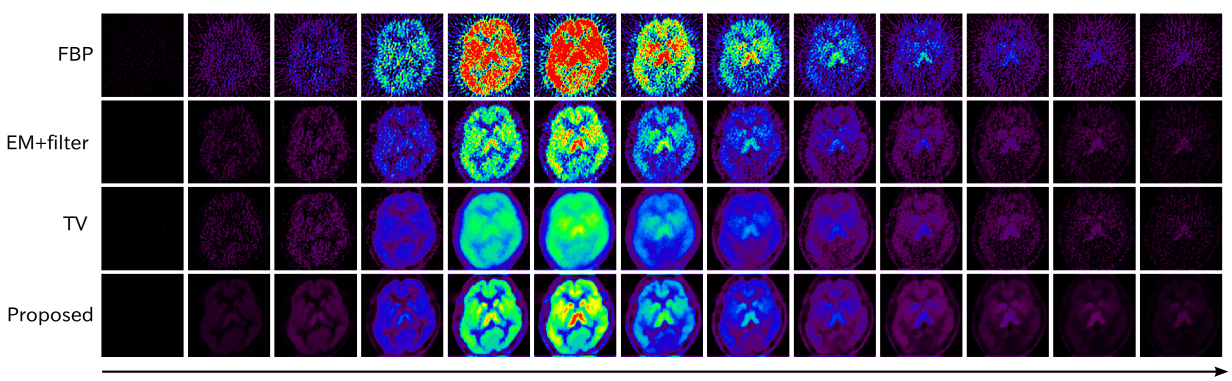


Fig. 7: Results of dynamic PET image reconstruction of clinical data: Subject 3.

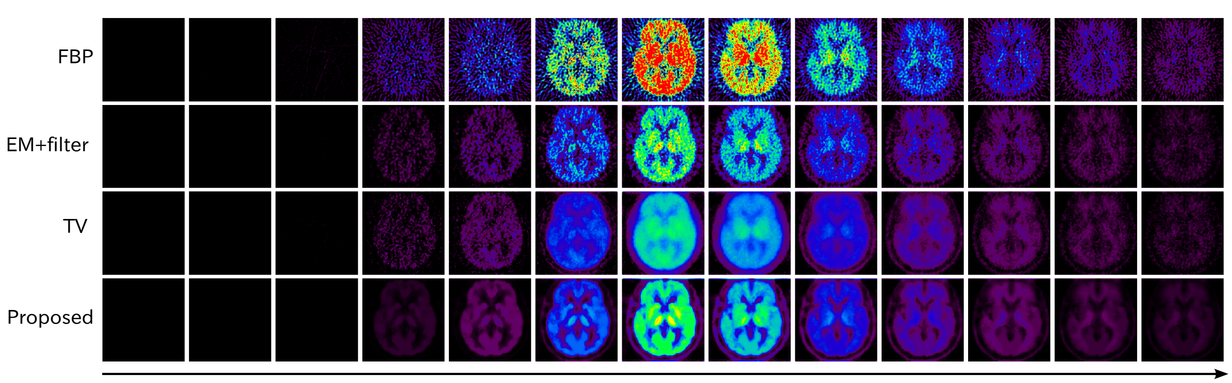


Fig. 8: Results of dynamic PET image reconstruction of clinical data: Subject 4.

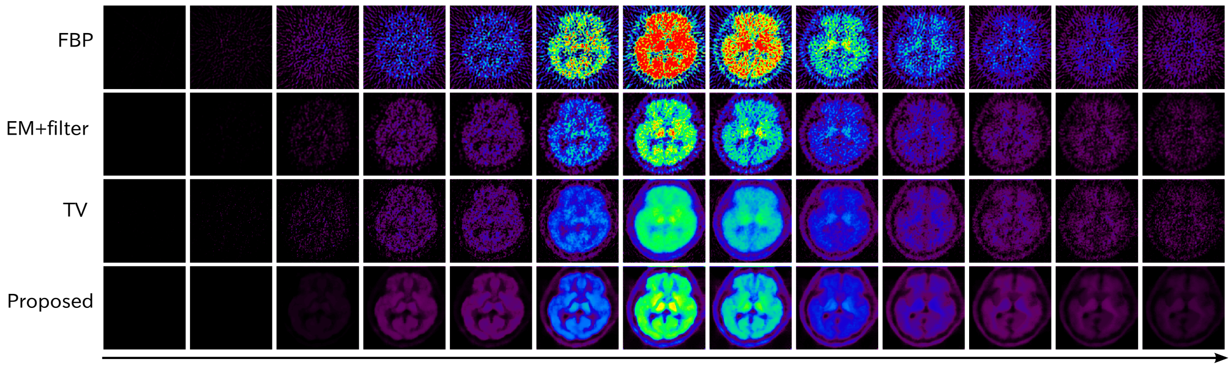


Fig. 9: Results of dynamic PET image reconstruction of clinical data: Subject 5.

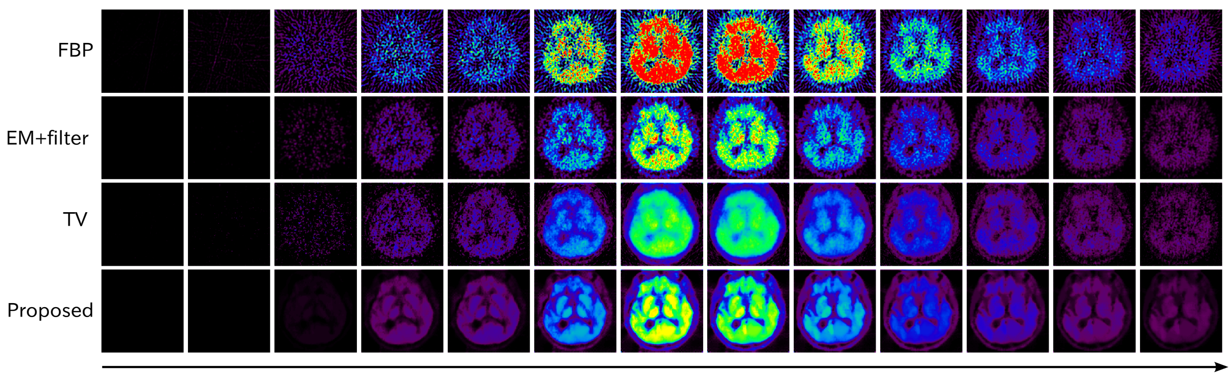


Fig. 10: Results of dynamic PET image reconstruction of clinical data: Subject 6.

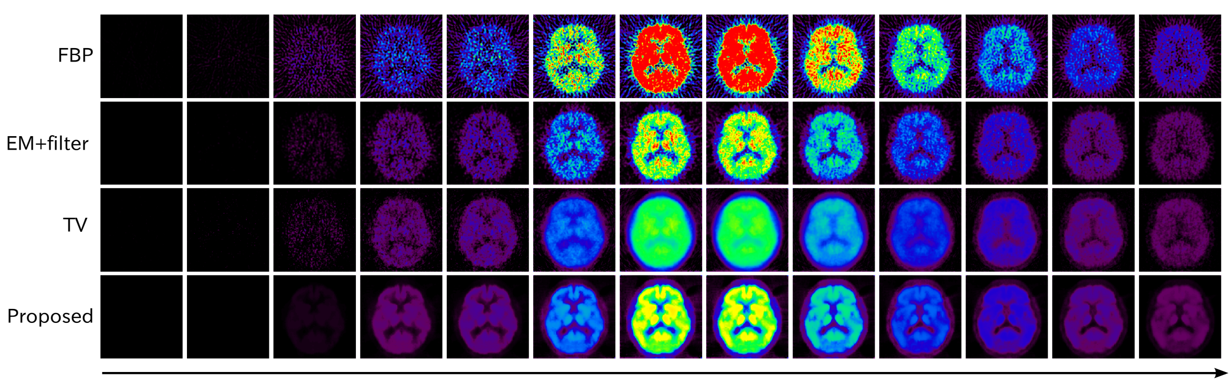


Fig. 11: Results of dynamic PET image reconstruction of clinical data: Subject 7.

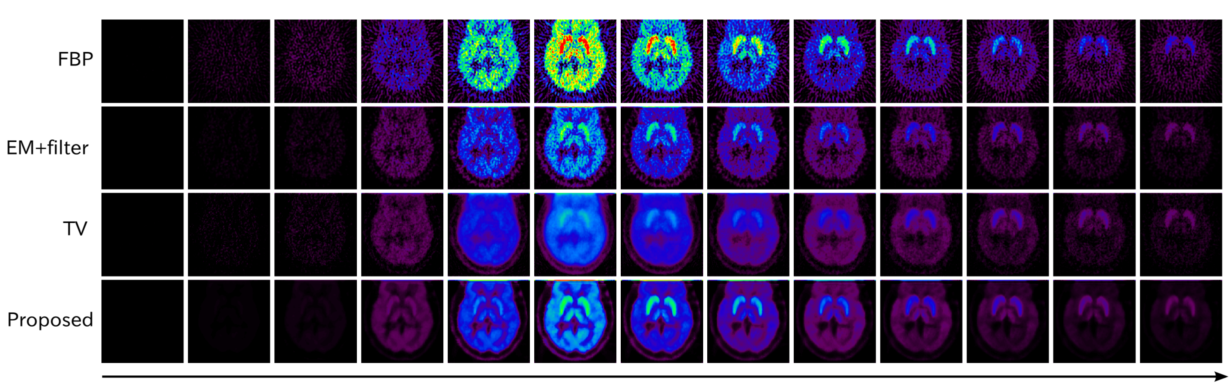


Fig. 12: Results of dynamic PET image reconstruction of clinical data: Subject 8.

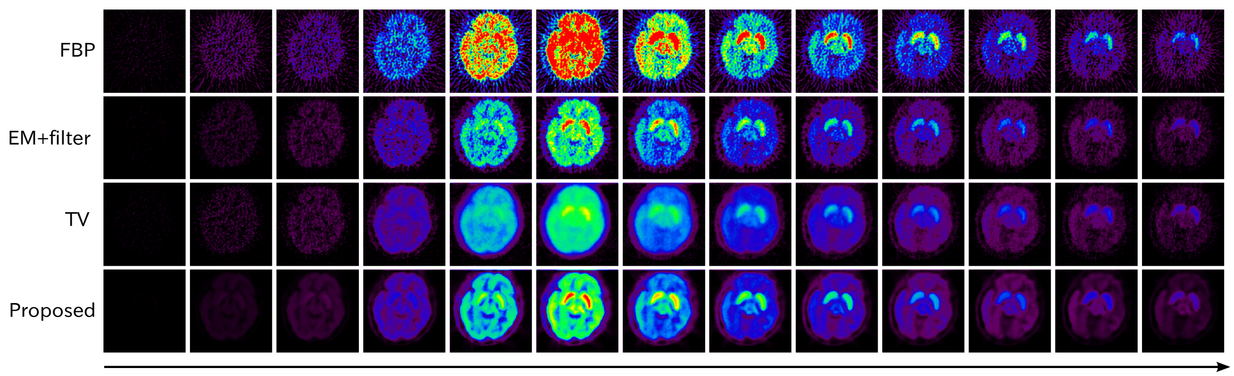


Fig. 13: Results of dynamic PET image reconstruction of clinical data: Subject 9.

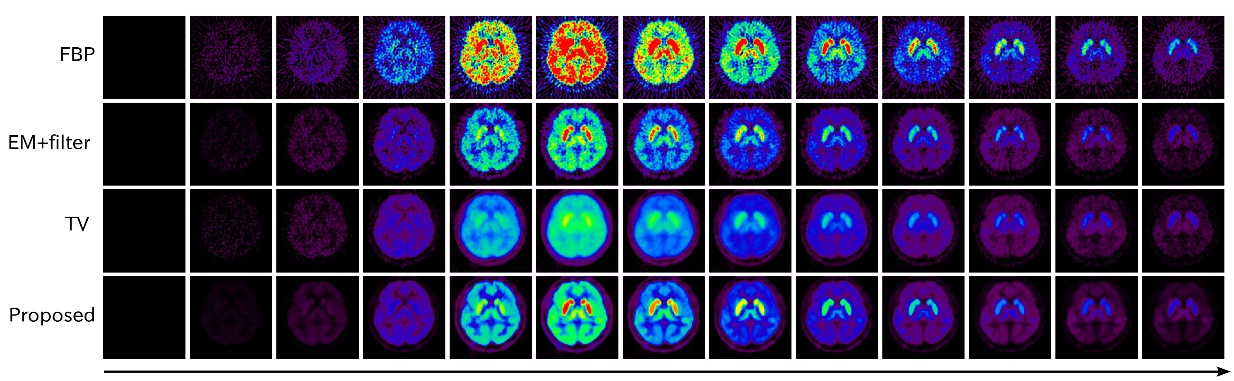


Fig. 14: Results of dynamic PET image reconstruction of clinical data: Subject 10.