Iterative PET Image Reconstruction Using Convolutional Neural Network Representation

Kuang Gong, Jiahui Guan, Kyungsang Kim, Xuezhu Zhang, Jaewon Yang, Youngho Seo, Georges El Fakhri, Jinyi Qi* and Quanzheng Li*

UC Davis, UC San Francisco,
Mass General Hospital, Harvard Medical School

IEEE TMI, Sept. 2018

Slides compiled by Lars Gjesteby

December 5, 2018

Motivation

- PET recon is an ill-posed inverse problem
- Low photon counts
- Denoising methods include Gaussian filtering and CNN post-processing
- Goal is to incorporate CNN into iterative recon pipeline – alternating direction method of multipliers (ADMM) algorithm

PET Data Model

In PET image reconstruction, the measured data $\boldsymbol{y} \in \mathbb{R}^{M\times 1}$ can be modeled as a collection of independent Poisson random variables and its mean $\bar{\boldsymbol{y}} \in \mathbb{R}^{M\times 1}$ is related to the unknown image $\boldsymbol{x} \in \mathbb{R}^{N\times 1}$ through an affine transform

$$\bar{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{s} + \boldsymbol{r}, \tag{1}$$

where $\boldsymbol{P} \in \mathbb{R}^{M\times N}$ is the detection probability matrix, with P_{ij} denoting the probability of photons originating from voxel j being detected by detector i [39]. $\boldsymbol{s} \in \mathbb{R}^{M\times 1}$ denotes the expectation of scattered events, and $\boldsymbol{r} \in \mathbb{R}^{M\times 1}$ denotes the expectation of random coincidences. M is the number of lines of response (LOR) and N is the number of pixels in image space. The log-likelihood function can be written as

$$L(\boldsymbol{y}|\boldsymbol{x}) = \sum_{i=1}^{M} y_i \log \bar{y}_i - \bar{y}_i - \log y_i! \,. \tag{2}$$

The maximum likelihood estimate of the unknown image $\boldsymbol{x}$ can be found by

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\geq 0} L(\boldsymbol{y}|\boldsymbol{x}). \tag{3}$$

Neural Network Representation

$$\boldsymbol{x} = f(\boldsymbol{\alpha}), \qquad (4)$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ represents the neural network and $\boldsymbol{\alpha}$ denotes the input to the neural network. Through this representation, inter-patient information and intra-patient information can be included into the iterative reconstruction framework through pre-training the neural network using existing data.

When substituting the representation in (4) using the above mentioned network structure, the original PET system model shown in (1) can be rewritten as

$$\bar{\boldsymbol{y}} = \boldsymbol{P} f(\boldsymbol{\alpha}) + \boldsymbol{s} + \boldsymbol{r}. \qquad (5)$$

The maximum likelihood estimate of the unknown image $\boldsymbol{x}$ can be calculated as

$$\hat{\boldsymbol{x}} = f(\hat{\boldsymbol{\alpha}}), \qquad (6)$$

$$\hat{\boldsymbol{\alpha}} = \arg\max_{\boldsymbol{\alpha} \geq 0} L(\boldsymbol{y}|\boldsymbol{\alpha}). \qquad (7)$$

The objective function in (7) is difficult to solve due to the nonlinearity of the neural network representation. Here we transfer it to the constrained format as below

$$\begin{aligned} \max \quad & L(\boldsymbol{y}|\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} = f(\boldsymbol{\alpha}). \end{aligned} \qquad (8)$$

We use the Augmented Lagrangian format for the constrained optimization problem in (8) as

$$L_\rho = L(\boldsymbol{y}|\boldsymbol{x}) - \frac{\rho}{2}\|\boldsymbol{x} - f(\boldsymbol{\alpha}) + \boldsymbol{\mu}\|^2 + \frac{\rho}{2}\|\boldsymbol{\mu}\|^2, \qquad (9)$$

Algorithm

Algorithm 1 Algorithm for iterative PET reconstruction incorporating convolutional neural network

Input: Maximum iteration number `MaxIt`, sub-iteration numbers `SubIt1` and `SubIt2`, penalty parameter ρ, image initialization $\boldsymbol{x}_{\text{ini}}$

Initialize $\boldsymbol{\alpha}^0 = f(\boldsymbol{x}_{\text{ini}})$, $\boldsymbol{x}^0 = f(\boldsymbol{x}_{\text{ini}})$, $\boldsymbol{\mu}^0 = \mathbf{0}$

for $n = 1$ to `MaxIt` **do**

  $\boldsymbol{x}^{n,0} = \boldsymbol{x}^{n-1}$

  for $t = 1$ to `SubIt1` **do**

    $\hat{x}_{j,\text{EM}}^{n,t} = x_j^{n,t-1}/p_j \sum_{i=1}^{n_i} p_{ij} \frac{y_i}{[\boldsymbol{P}\boldsymbol{x}^{n,t-1}]_i + s_i + r_i}$, where $p_j = \sum_{i=1}^{n_i} p_{ij}$

    $x_j^{n,t} = \frac{1}{2}\Big[f(\boldsymbol{\alpha}^{n-1})_j - \mu_j^{n-1} - p_j/\rho + \sqrt{[f(\boldsymbol{\alpha}^{n-1})_j - \mu_j^{n-1} - p_j/\rho]^2 + 4\hat{x}_{j,\text{EM}}^{n,t} p_j/\rho}\Big]$

  end for

  $\boldsymbol{x}^n = \boldsymbol{x}^{n,\texttt{SubIt1}}$, $\boldsymbol{\alpha}^{n,0} = \boldsymbol{\alpha}^{n-1}$, $\boldsymbol{\theta}^{n,0} = \boldsymbol{\alpha}^{n-1}$, $t_0 = 1$

  for $k = 1$ to `SubIt2` **do**

    $t_k = (1 + \sqrt{1 + 4t_{k-1}^2})/2$

    $\alpha_j^{n,k} = \Big[\theta_j^{n,k-1} - \beta \sum_{m=1}^{N} \frac{\partial f(\boldsymbol{\theta}^{n,k-1})_m}{\partial \theta_j} [f(\boldsymbol{\theta}^{n,k-1})_m - x_m^n - \mu_m^{n-1}]\Big]_+$

    $\boldsymbol{\theta}^{n,k} = \boldsymbol{\alpha}^{n,k} + \frac{t_{k-1}-1}{t_k}(\boldsymbol{\alpha}^{n,k} - \boldsymbol{\alpha}^{n,k-1})$

  end for

  $\boldsymbol{\alpha}^n = \boldsymbol{\alpha}^{n,\texttt{SubIt2}}$

  $\boldsymbol{\mu}^n = \boldsymbol{\mu}^{n-1} + \boldsymbol{x}^n - f(\boldsymbol{\alpha}^n)$

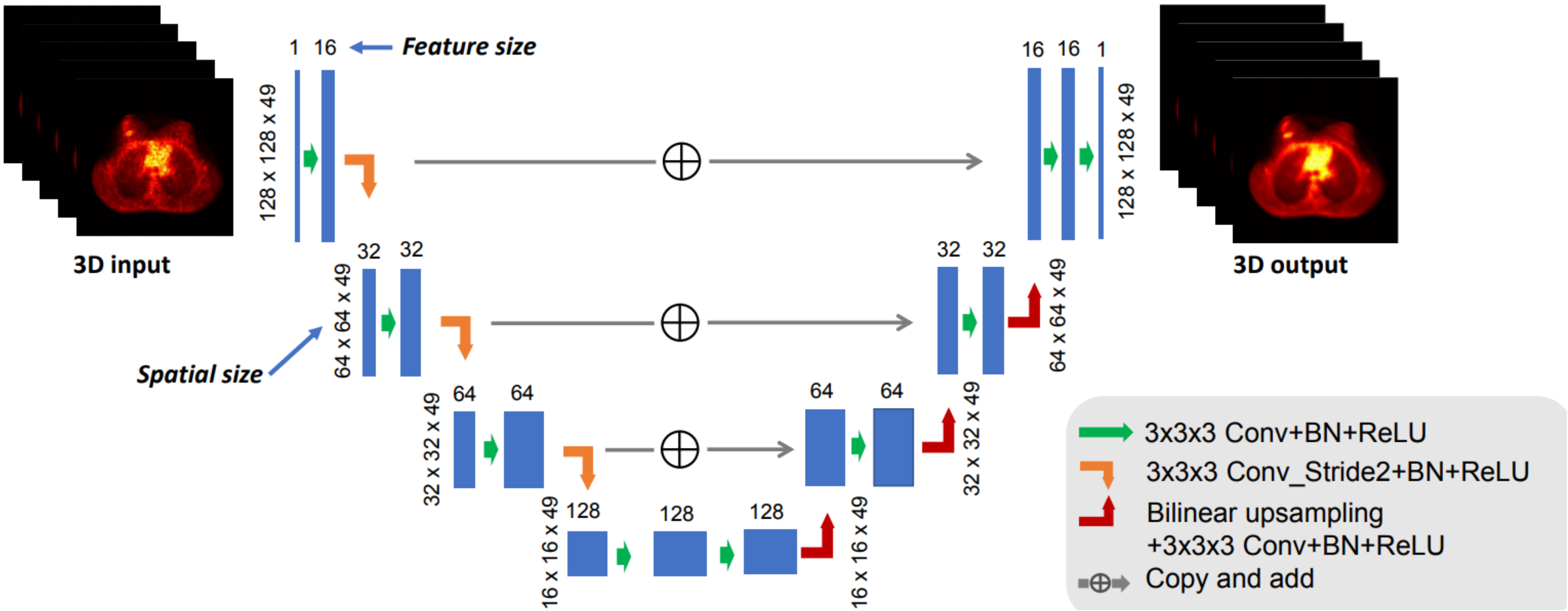
end for

return $\hat{\boldsymbol{x}} = f(\boldsymbol{\alpha}^{\texttt{MaxIt}})$

Approach

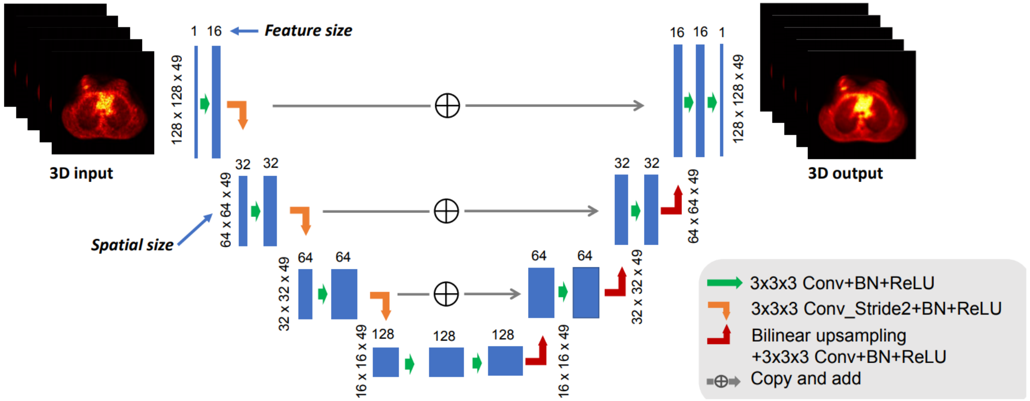
- Modified U-net and ResNet structure
- 3D convolutions
- Dynamic inter-patient and intra-patient data
- CNN Input: iterative reconstructions of low-counts data
- CNN Label: high-counts reconstructions to represent the unknown PET image to be reconstructed
- CNN defines the feasible set of valid PET images

Network Structure



Network Modifications

- Conv layer with stride 2 to down-sample image instead of max pooling



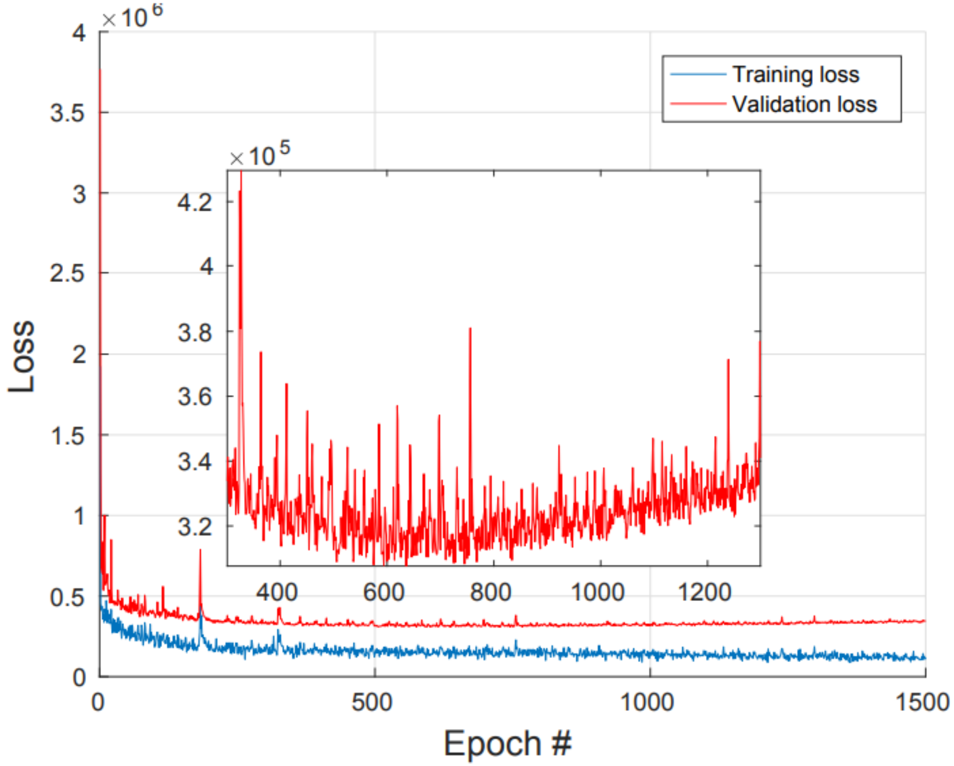
- Bilinear interpolation for upsampling instead of transpose convolution to overcome checkerboard artifacts
- Adding left side feature to right side to reduce training parameters

Datasets

- Simulated data from XCAT phantoms
  - 51 3D datasets containing 49 axial slices of organ activity regions
- Lung patient study
  - 6 patients, 45 3D datasets containing 49 slices
- Brain patient study
  - 17 patients, 45 3D datasets containing 91 slices

Training

- Adam optimization
- L2 norm
- 600 epochs for simulation study



Results: Simulation Study

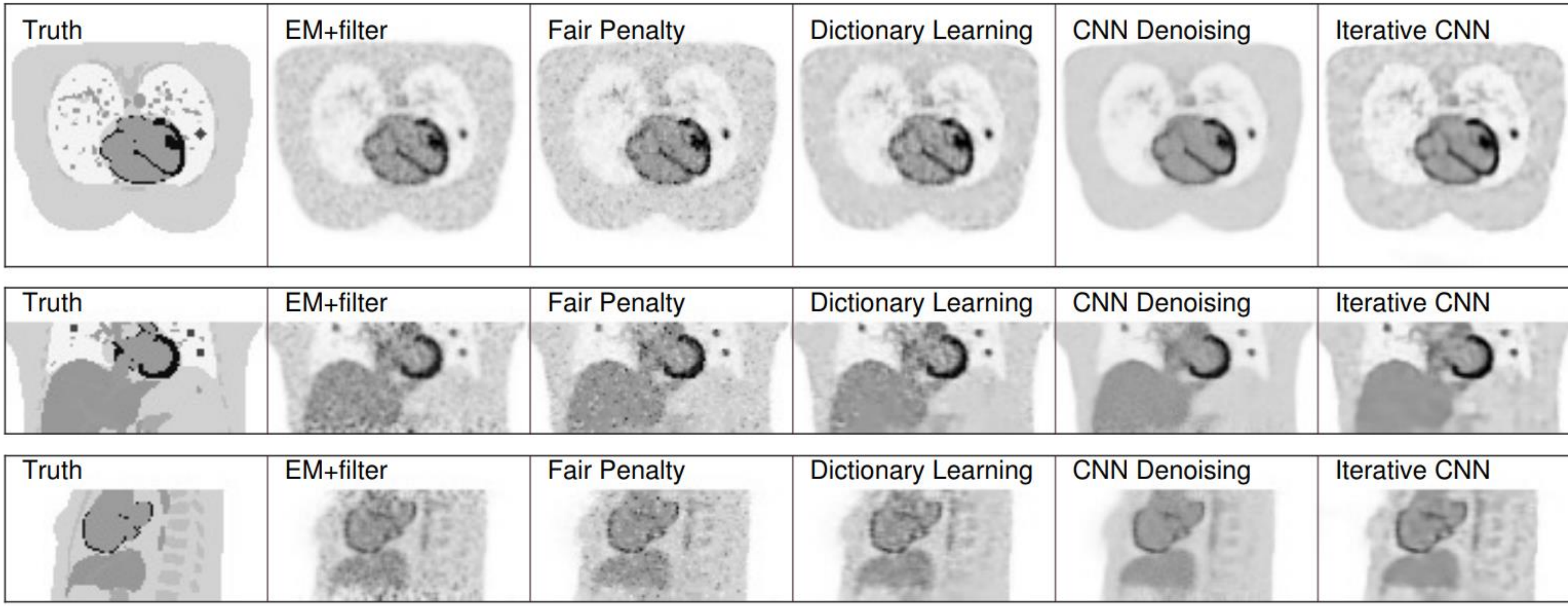


Fig. 6: Three views of the reconstructed images using different methods for the simulation data set. From left to right: ground truth, Gaussian denoising,fair penalty based penalized reconstruction, dictionary learning based reconstruction, CNN denoising, and the proposed iterative CNN reconstruction

Results: Simulation Study

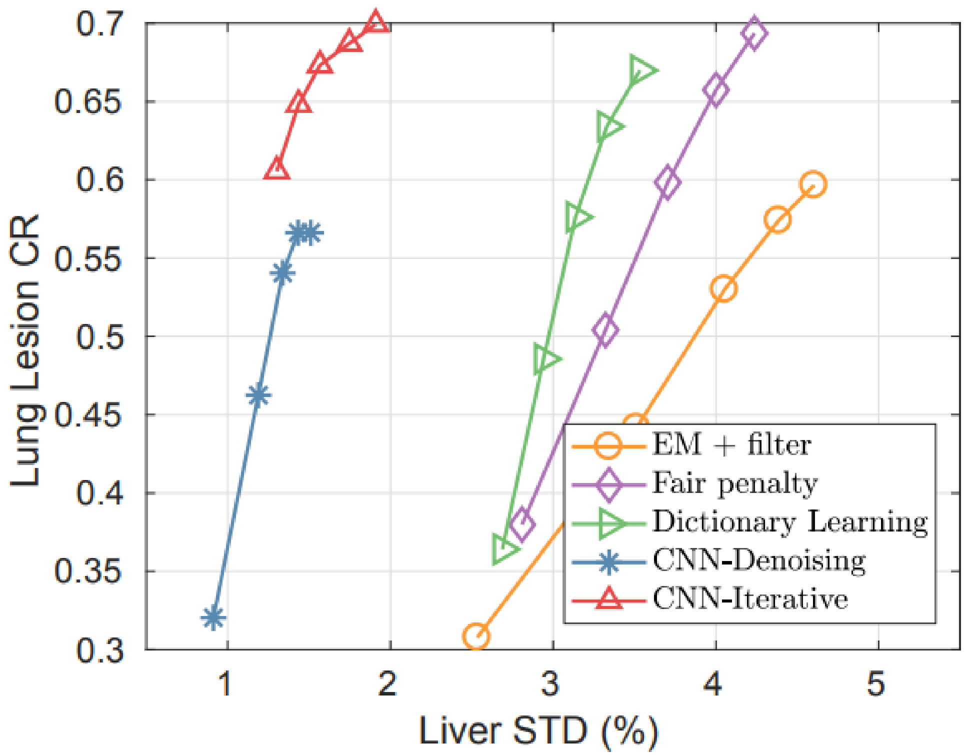


Fig. 7: The contrast recovery vs. STD curves using different methods for the simulated data sets. Markers are plotted every twenty iterations with the lowest point corresponding to the 20th iteration.

Results: Lung Study

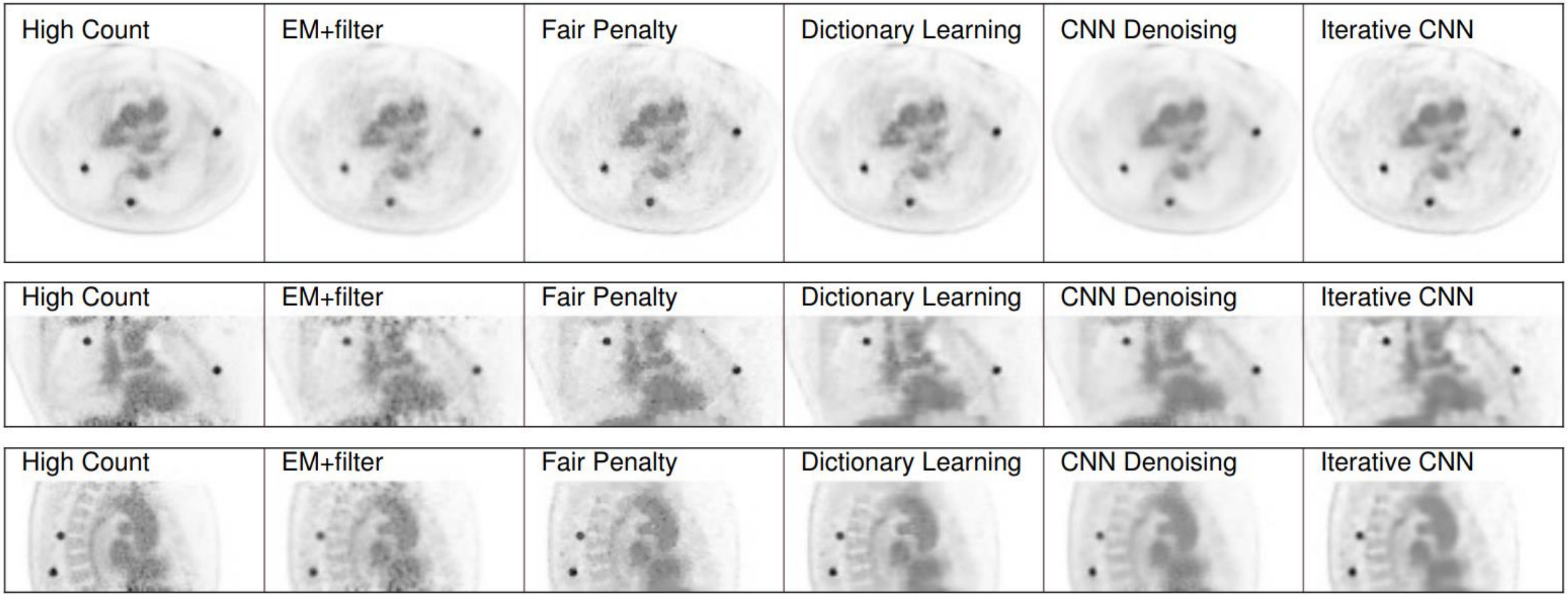


Fig. 8: Three views of the reconstructed images using different methods for the lung patient set. From left to right: high count reference image, Gaussian denoising,fair penalty based penalized reconstruction, dictionary learning based reconstruction, CNN denoising, and the proposed iterative CNN reconstruction

Results: Lung Study

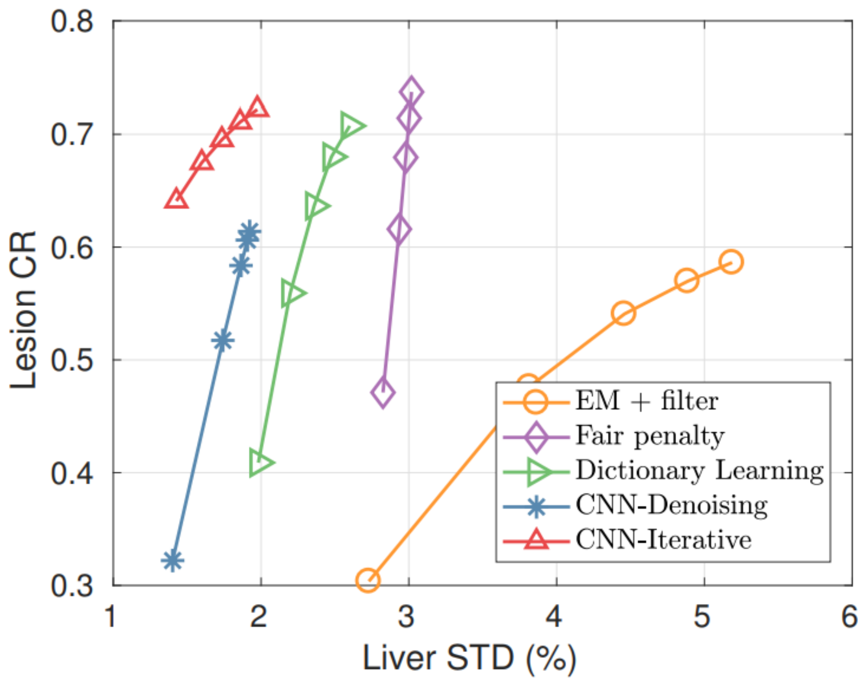


Fig. 9: The contrast recovery vs. STD curves using different methods for the real lung data sets. Markers are plotted every twenty iterations with the lowest point corresponding to the 20th iteration.

Results: Brain Study

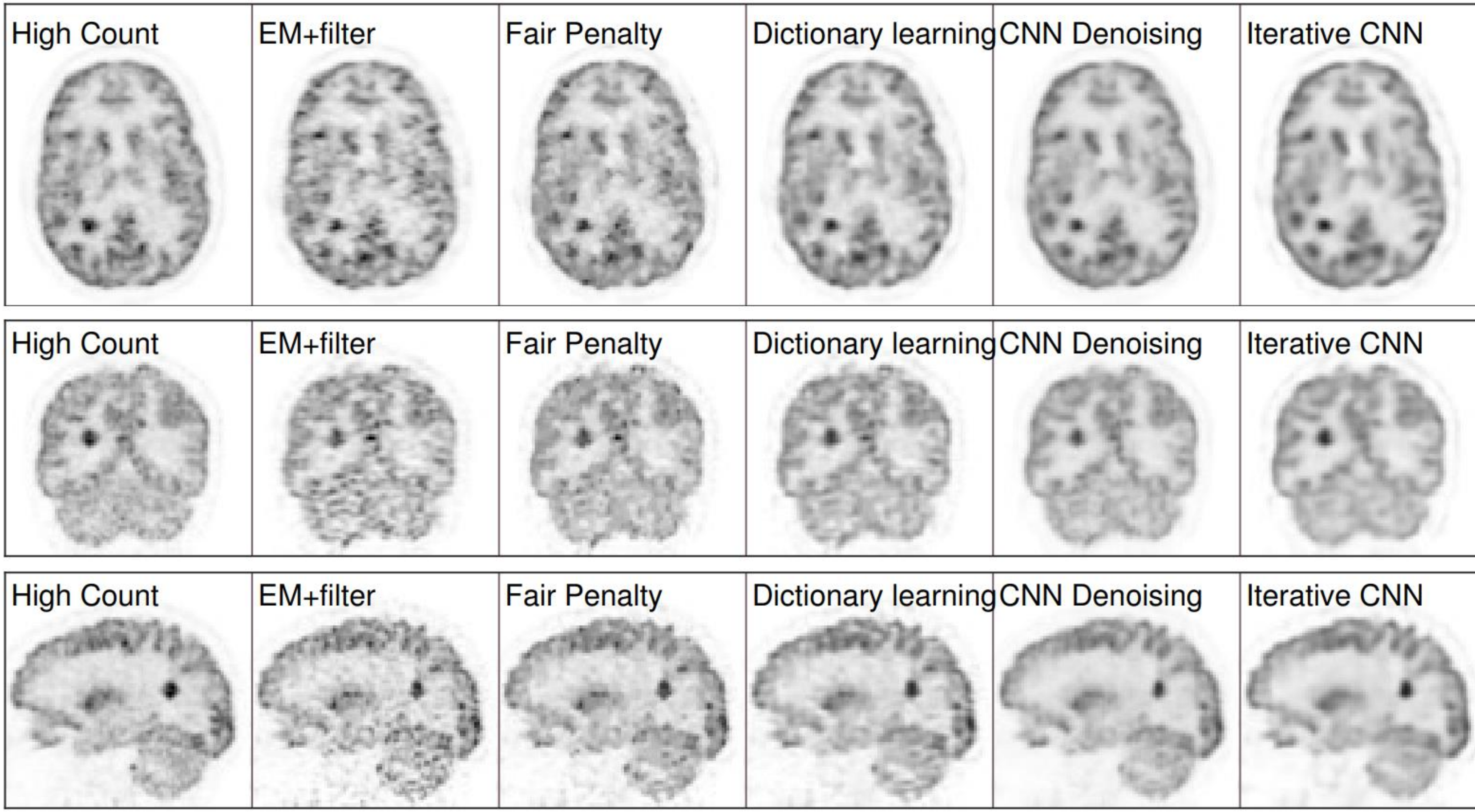


Fig. 10: Three views of the reconstructed images using different methods for the brain patient data. From left to right: high count reference image, Gaussian denoising,fair penalty based penalized reconstruction, dictionary learning based reconstruction, CNN denoising, and the proposed iterative CNN reconstruction.

Results: Brain Study

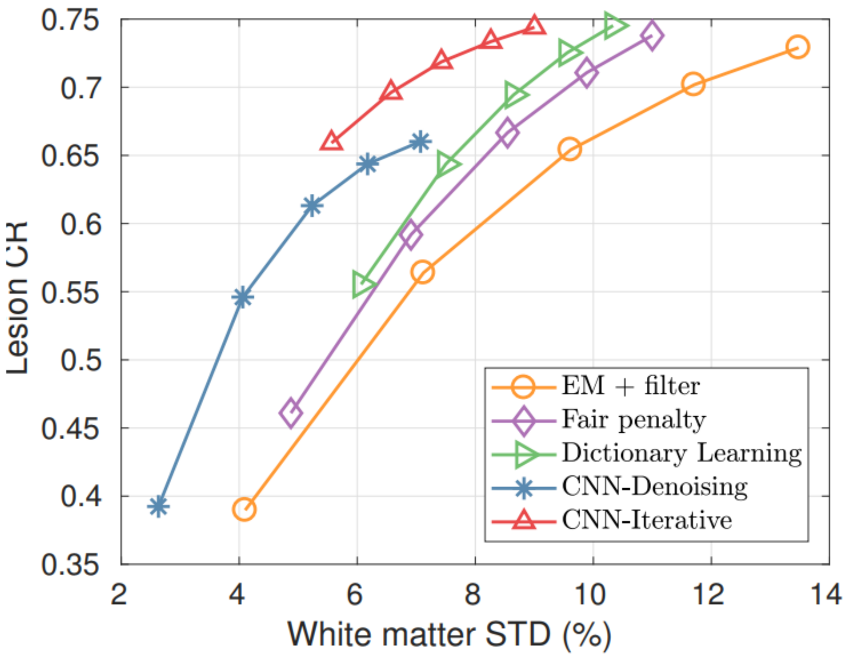


Fig. 11: The contrast recovery vs. STD curves using different methods for the brain data sets. Markers are plotted every twenty iterations with the lowest point corresponding to the 20th iteration.

Discussion

- CNN for image representation in iterative recon framework outperforms existing denoising methods
- CNN is data-driven and can accommodate prior information from multiple sources (temporal, anatomical, inter-patient)
- Training data without lesions led to poor lesion recovery in CNN validation sets
- Small structures are lost in brain study results