PET Image Reconstruction using Convolutional Neural Network and Generative Adversarial Network in Sinogram Domain

IGNAS KUPCIKEVICIUS

Biomedical Engineering & Informatics
School of Medicine and Health, Aalborg Univertity

June 2019

Abstract

Introduction: *Artifacts and noise in PET imaging are caused by multiple factors, including lowered radiotracer dose and information loss in the form of missing pixels or missing projections. Recently, deep learning based algorithms have achieved promising results in the medical imaging field, including PET denoising and CT sinogram inpainting, especially using the Convolutional Neural Network (CNN) and Generative Adversarial Networks (GAN) architectures.* ***Aim:*** *This article aims to compare CNN and GAN approaches for missing data reconstruction on PET sinogram domain.* ***Methods:*** *The end-to-end framework, from PET image to the sinogram domain and back to PET image domain, was proposed. The Radon transform was applied to covert PET images into sinograms. The first model was the CNN encoder-decoder based network with four skip connections. The effective strategy was applied to efficiently train more corrupted PET sinograms by loading previously trained weights. The second approach was the GAN network, with the generator designed similarly as the CNN encoder-decoder, and the discriminator containing four convolutional layers to classify generated sinograms as artificially generated or ground truth. The proposed framework ended by applying filter back projection algorithm to transform sinograms back to PET image domain.* ***Results:*** *The results revealed that GAN outperformed CNN by a small margin. The average PSNR and SSIM scores within all five corruption levels were 41.44, 0.977, and 42.34, 0.983 when predicting missing pixels. Differences of two metrics between CNN and GAN were higher when predicting missing projections; 40.13, 0.866 versus 46.84, 0.989. Additionally, GAN performed noticeably better when 90% of sinogram data were removed, resulting in a sharper and more detailed reconstructed image, qualitatively comparing to CNN.* ***Discussion:*** *Different network architectures, chosen hyperparameters, and objective functions, might be the reasons why GAN performed better than CNN. Even though the study had some limitations, the promising results were achieved, which motivates to experiment further.*

Keywords PET reconstruction, sinogram domain, CNN, GAN, FBP, Radon transform

1. INTRODUCTION

Positron Emission Tomography (PET) is one of the major imaging modalities, broadly used in hospitals for diagnostic and intervention purposes [1]. High-quality PET images are often used in diagnosing brain diseases and disorders, such as Alzheimer's, dementia [2]. Unfortunately, reconstructed PET images usually contain more noise and artifacts compared to magnetic resonance imaging (MRI) and computed tomography (CT) due to the constraints, such as low number of coincidence events - photons that reach the detector, imposed by a reduced dose of a radiotracer.[3] Additionally, the noise and artifacts in PET are caused by the missing pixels or missing projections which appear in high-resolution PET scanners [4, 5]. All these factors contribute to PET image quality by making the resulting image blurred, unclear, and hard to interpret. Therefore, scientists and researchers seek to come up with different solutions to overcome the limitations and improve PET image quality. Additionally, the ability to use fewer data and predict high-quality medical images would require less radiation to achieve the desired outcome. There is a trade-off between high-quality PET images and restrictions of available events from minimizing exposure for health reasons [6]. If high-quality

PET images could be predicted from lower data representations, a lower amount of radiotracer would need to be used. PET system detects pairs of gamma rays emitted from the injected radioactive tracers on the biologically active molecule. These coincidental gamma pairs are recorded to generate a 3D volume using computer analysis [2]. According to the report "Biological Effects of Ionizing Radiation (BEIR VII)", the estimated risk of cancer increases about 0.04% per single PET scan, which multiplies after repetitive examinations. On the other hand, when reducing the radiation dose, the quality of PET images are degraded, due to the lower amount of detected photons [1], which negatively affects the diagnostic capabilities of PET. To address radiotracer dose related issues, Shi et al. [1] used regression forest-based framework and predicted normal dose PET images from low-dose images. The study showed promising results when the mean standardized uptake value of predicted standard dose images was close to the ground truth and proved that the machine learning approach could be successfully used on PET images. However, it also had limitations when images were predicted voxelwise and did not account for the relationship with the neighboring voxels. Moreover, the study was conducted using a small number of subjects.

Recently, PET systems started using silicon photomultipliers (SiPMs) as a detector for photon registration. According to the literature [5], there is a high probability that some of the SiPMSs may not function properly, resulting in loss of counts in pixels. Such "dead" pixels are pixels with missing data values causing the degraded quality of reconstructed PET images. Additionally, the missing pixels may arise due to the problems in processing electronics [5]. The study by Jong et al. [4] investigated a few methods to compensate for missing data caused by the gaps between high-resolution PET scanner detectors and faulty electric parts. Two approached were applied: bilinear interpolation approach and model-based approach. The model-based approach outperformed bilinear interpolation, but with a price of long processing time. [4]

A variety of studies used commonly know algorithms for improving PET image quality. Firstly, Gaussian filtering and bilateral filtering were used to improve PET post-reconstruction; however, such filtering resulted in reduced edge preservation and blurred images [7]. Furthermore, the block-matching 3D (BM3D) [8] approach was applied to PET images and showed structural improvements [9]. However, the study used simulated, not the real PET data, while BM3D is not effective on natural images due to the lack of highly correlated matching blocks.[10] Another attempt to improve medical imaging quality is to apply filtering technique on the sinogram domain instead of reconstructed images. A sinogram consists of an angle based histograms. Such data format is usually the raw format in PET systems.[2] In theory, when saving PET data in the sinogram domain, every detected event is stored using the angle and the offset of the scanned tissue point, which represents a single event. Multiple sorted events create the sinogram (Figure 1), which usually is a smaller data format compared to the reconstructed image.

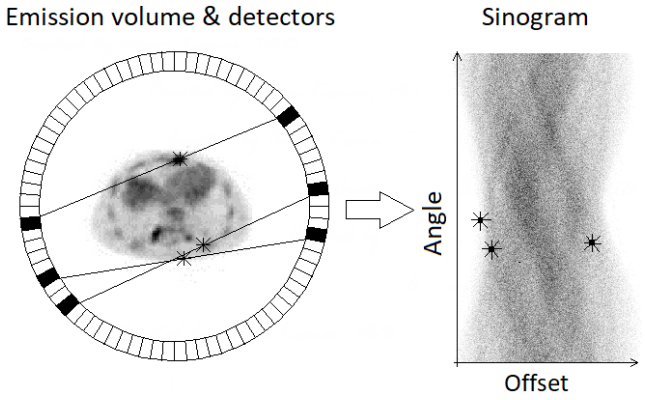

Figure 1: *Photon annihilation events in PET detector ring.*

A comparison study by Yu and Muhammed [11] suggested that signal-to-noise ratio (SNR) was improved on pre-reconstruction denoising, where PET sinogram are denoised before applying a reconstruction algorithm, compared to post-reconstruction algorithms. However, the study used only the classical filtering approaches which tend to induce artifacts; the study was not extended to investigate the performance with more recent neural networks based approaches.

Finally, most recently, convolutions neural networks (CNN) has become a state-of-the-art technique in terms of the image analysis [12]. A study by An et al. [13] proposed a deep CNN for full-dose PET image reconstruction based on the local patches from the low-dose PET. CNN was applied to the CT sinograms to predict the residuals - a difference between noise

and ground truth images. Residuals were subtracted from the noisy sinograms, resulting in a higher quality of the reconstructed CT images [12]. Even though the study showed promising results on CT, it is still unknown how sinogram based reconstruction using CNN would perform on PET images.
Another widespread machine learning approach is Generative Adversarial Network (GAN) recently received widespread attention [14, 15, 6]. GANs are known as generative models, which have two main components: generator and discriminator. The generator learns to replicate the realistic input images, while the discriminator tries to distinguish between the generated and real images. A study by Zhao et al. [16] suggested a sinogram inpainting network which uses GAN to solve a limited angle CT reconstruction problem. The study concluded promising results in CT sinogram inpainting, yet again, there were no indications of how such an approach can be applied to PET images. Therefore, it is still unknown how CNN and GAN would perform on PET sinograms. This is the key interest of this study.

This study aims to use CNN and GAN approaches for corrupted PET image reconstruction in the sinogram domain. Additionally, to observe the dependency of models performance to the available image data, two data representations of multiple levels of corruption are introduced.

2. METHODS

Data acquisition

The publicly available dataset was acquired from the Cancer Imaging Archive (TCIA: http://www.cancerimagingarchive.net), an archive of the medical images of cancer created by the National Cancer Institute (NCI) [17]. Soft-tissue-Sarcoma dataset was chosen from TCIA, which is, based on the thorough search of the relevant datasets yielded a large number of tested subjects, giving the high number of training samples, necessary to train a deep neural network [18]. The dataset consists of 51 patients with histologically proven soft-tissue-sarcomas of the lower extremities in PET/CT/MRI image formats [17, 19].

For this study PET data from all 51 subjects were used which contained DICOM (Digital Imaging and Communications in Medicine) images of 128 x 128-pixel resolution in a total of 13,417 2D images.

The fludeoxyglucose (FDG) radiotracer was used when performing PET scans on a PET/CT scanner (Discovery ST, by GE Healthcare). A median of 420 MBq of FDG was injected intravenously followed by the 60 min of body imaging acquisition.[17] Additionally, the attenuation correction was performed which reduces severe artifacts induced due to the high number of lost photons during the procedure.[17]
From 51 data subjects, five random patients were put aside and left as the unseen test dataset, which approximately corresponds to the 10% of the entire dataset. This test set is used only for the final results. The following 46 subjects correspond to the 12,082 PET images, which were divided into training and validation datasets with a ratio of 70% to 30%, respectively. Training and validation data sets were used during the training phase and for hyperparameter tuning. After each training phase, the qualitative results were observed for evaluation of reconstructed images, in terms of visible structural details, artifacts, blur. Qualitative results represent the real-world scenario since this is how the doctors perceived image information in the working environment. Any image distortions increase the risk of misdiagnoses. Thus, the quality of reconstructed images have a direct impact on patient treatment [20].

Radon transform

The TCIA data was in reconstructed PET image format. Therefore, the Radon transform was applied to convert data into the raw sinogram domain. Radon transform of the image described as the function $f(x, y)$, is defined by the combination of the integrals through $f(x, y)$ in vertical and horizontal projections, or y and x axis, respectively. [21] Radon transform mathematical is defined as:

$$Rf = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(xcos\theta + ysin\theta - s)dxdy$$

where θ is the angle of the line and s is a perpendicular offset of the line. [21]

In this study, the data were transformed by using *skimage* library in Python coding environment. *Skimage* allowed to define two parameters before

transformation: the angle *theta*, which was left as a default value of 180°, and *Circle*, which when set on 'True' allows making a uniform sinogram size of 128 projections and 128 detectors, corresponding to the size as the original image. This property was helpful since CNN requires uniform dataset (Figure 2).

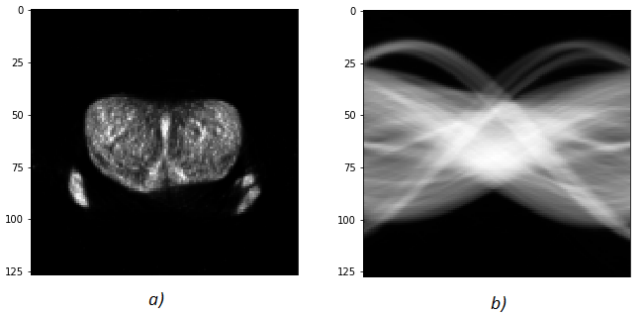

Figure 2: *Radon transform applied on the PET image (a). Resulting sinogram (b) is used for further pre-processing.*

Two data representations were created to address missing data problems in PET imaging domain: missing pixels and missing projections. Missing pixels are simulated by removing random data points from the PET sinogram. The amount of randomness was increased in 5 levels, 10%, 30%, 50%, 70% and 90% (Figure 3 a)), to reflect multiple scenarios of missing patterns, also, to investigate how far CNN and GAN can perform to compensate missing pixel values properly. The second data representation was focused on simulating missing projections, which appear in PET sinograms due to the gaps between detector heads in the scanner or because of faulty detectors [4]. Also presented in 5 different levels (Figure 3 b))

Intensity normalisation

Applying the Radon transform converted the whole data into the sinogram domain. Therefore, further pre-processing was applied to prepare data for the neural networks. Data were rescaled so that all the values would be within the range 0 and 1. Such intensity normalization helped to maintain the weights as relatively small values during the training phase. The large weight values usually result in the poor performance and unstable model, thus, it is critical to rescale input and output data before presenting data to the network. [22] Thereby, this was achieved by applying the following equation:

$$Z_i = \frac{Y_i - Y_{min}}{Y_{max} - Y_{min}}$$

where Z_i is rescaled data, Y_i original data.

Model architecture

In this paper, two different architectures were used and compared for lost data reconstruction in PET sinogram domain. Most of the hyperparameter tuning was performed manually, where some of them were chosen according to the literature. Rectified linear unit (ReLu) was used as an activation function throughout the proposed networks since it mathematically simple, yet effective, also recommended according to the literature [18]. Different optimizers were tested, and Adam solver [23], with a learning rate of 0,003 showed the fastest convergence and best results and both CNN and GAN models.

CNN Architecture

The first architecture is a convolutional neural network based autoencoder inspired by U-net CNN for image segmentation [24]. Proposed network has four convolutional layers with 32, 48, 64, 64 filters, respectively. These layers encoded the image data into a smaller form, which contained automatically learned the most important features. Then, the compressed data were fed through four deconvolving layers of 64, 64, 48, 32 filters, respectively. Deconvolutional layers of the network are called a decoder part. It reconstructed the compressed image data into the normal size (128 x 128) images. 3 x 3 filter size was used through the entire network, with a stride of one, based on the similar studies [24, 25]. Additionally, four skip connections were added which shuttled low-level features to the high level features [16]. Such direct connections were proven to strengthen feature propagation, encourage feature reuse, and lower the computational cost by reducing the number of parameters.[26] In the CNN model, the training and validation losses after every batch were computed between the batch of predicted and the ground truth values using Mean-square-error (MSE) also known as L2 loss. MSE is one of the commonly used objective function, and it is available by default on Keras library. It was used in the CNN model since it is a *de*

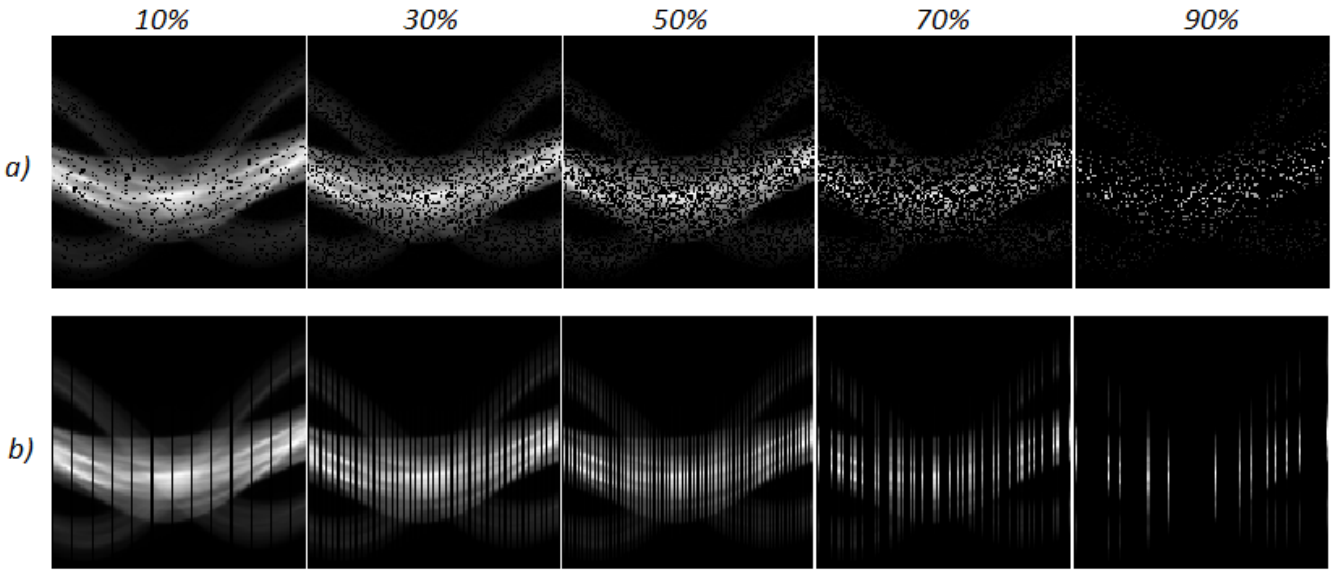

Figure 3: *Representations of corrupted sinograms with (a) simulated missing pixels, and (b) missing projections at 5 different levels.*

facto standard objective function in neural networks [27].

Two CNN models were trained from scratch; one with data representation of 10% of missing pixels, and the second with a similar amount of missing projections. Models were trained for 100 epochs until MSE error stopped decreasing, which indicated that the models started to overfit. The weights from these models were loaded to retrain the data representations with a higher level of corruption. This time approximately 30 epochs per all executed training were enough until the models started to overfit. Thus, using previously trained weights allowed to accelerate training by achieving faster convergence but maintaining the high quality of reconstructed PET sinograms. A similar strategy was reported when pretrained 2D weights were used for 3D model training [28].

GAN Architecture

In this paper, the proposed GAN was based on the 'pix2pix' network (Figure 4) [29], since it was previously successfully applied for the image inpainting problem when a big portion of removed sinogram was reconstructed with promising results [15]. Unlike conventional neural networks, GAN consist of two models: generator and discriminator. The GAN automatically learns the goal specific loss function, which classifies the output into real (ground truth) of 'fake' (artificially generated), while at the same time trains the generative model to minimize the loss error Isola et al. [29]. Blurred images are ranked as 'fake', therefore, the GAN will adjust generative model weights to generate more realistic PET images. Such GAN working principle is the biggest advantage against regular CNN and its pre-defined loss functions, such as mean-square-error (MSE) [28]. For GAN generator model, the previously described CNN encoder-decoder architecture with skip connections was used. The discriminator consists of 4 convolutional layers with 64, 128, 256, 512 filters, respectively. The final output layer was a convolutional layer with one filter, followed by a Sigmoid function.

In the GAN model, the objective function was similar to the one used in 'pix2pix' paper [29], and mathematically described as a combination of GAN objective and L1 loss, also known as mean-absolute-error:

$$G^* = arg \min_G \max_D \mathcal{L}_{GAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

where G^* is a final objective, G a generator, D a discriminator, L_{GAN} a Laplace transform of G,D objectives, L_{L1} a Laplace transform of L1 loss. G tries to minimize the objective function against the adversarial D that acts opposite and maximizes it [29].

Additionally, the discriminator was designed to pe-

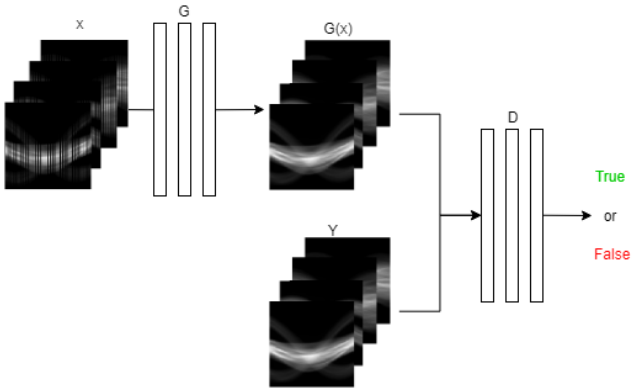

Figure 4: *Training a GAN by passing corrupted images (x) through the generator (G) to synthesize indistinguishable images (G_x) from the real images (Y). The discriminator (D) learns to classify between real and fake images, while generator tries to fool discriminator*

nalize prediction on the defined 4 x 4 pixels patch sizes instead of the full image. Such a strategy was suggested in the 'pix2pix' paper [29], which demonstrated that patches, smaller than the full-size image, produced a high-quality outcome, had fewer parameters and made the training faster.

Inverse Radon Transform

One of the easiest ways to reconstruct the image from the sinogram domain is by using filtered backprojection (FBP). When applying backprojection, the blurring effect was induced on image space. The additional filtering was used to overcome this limitation and to correct the blurring effect to some extent. [30] In this study, different available filters were tested to see which filter performed best for TCIA dataset. The ramp filter showed noticeably better results and was used when transforming the predicted test date data back to the image space.

Evaluation and validation

Observing image quality might be very subjective evaluation, therefore, the quantitative metrics were introduced. A commonly used metric within different imaging studies is the peak signal-to-noise ratio (PSNR).[25, 28] PSNR is the ratio measure between the power of image peak value and the power of the corruption that affects its representation, measured in decibels (dB) [10]. PSNR can be defined using the mean square error (MSE) measured between ground truth and corrupted representations:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

where m, n are image dimensions, I for monochrome image and K for its corrupted approximation, i, j coordination system values. Knowing that the signals might have a wide dynamic range, PSNR could be expressed in the logarithmic decibels (in dB):

$$PSNR = 10 \cdot log_{10}(\frac{DR^2}{MSE})$$

where DR is a dynamic range of the image. Usually, the greater value of PSNR means a higher quality reconstruction of the image. [25]

Additionally, structural similarity index measure (SSIM) metric Wang et al. [31] was used since PSNR does not guarantee the best perceptual and textural outcome, even when resulting score is high. PSNR metric measures per-pixel intensities, which is not a way of how a human perceives the image information. Additionally, according to the literature, this metric suffers from the regression-to-mean problem resulting in blurred reconstructed images. SSIM considers image corruption level as perceived change in image structural information measured between two images. SSIM compares groups of pixels between two images, using three composite measure: luminance, contrast, and structural. Simplified SSIM mathematical expression can be seen as:

$$SSIM(A, B) = l(A, B) \cdot c(A, B) \cdot s(A, B)$$

where A, B are groups of pixels from two images, l - luminance measure, c - contrast measure and s - structural measure. Since SSIM compares pixel patches, instead of individual pixels, this metric represents results more similar to how the humans observe image data. [31] Qualitative and quantitative results are evaluated using a test set of 1335 PET sinograms.

3. Results

Averaged PSNR, SSIM, and MSE values for different corruption levels and for both data representations

using CNN and GAN can be seen in (Figure 5). The top left graph indicated the continuous decline of PSNR value when more image information was excluded in data representations with removed pixels. PSNR at the last level of corruption was 35.43 for CNN and 36.91 for GAN. Interestingly, with data representation of missing projections (bottom left graph), the PSRN rapidly decayed only after removing more than 50% of the data points. The middle graphs of the SSIM metric demonstrated that both CNN and GAN performed well with values close to the perfect SSIM of 1 on data representation with missing pixels. The noticeable decline in SSIM was visible at higher corruption levels (70,90%) when removing data projections, but only with the CNN model. On the other hand, GAN preserved the high quality of reconstructed images and showed stable performance throughout all levels of corruption. The right two graphs identify how the MSE increases when the number of lost information increases.
All three metrics revealed that GAN network outperformed CNN autoencoder throughout the all corruption level of both data representations. Exceptionally, PSNR was higher of CNN network compared to GAN only on data representation with 70% missing pixels, with 40,1 for CNN and 39,98 for GAN, respectively.

Qualitatively results (Figure 6) of the sample image from the test set indicated how well both CNN and GAN models managed to reconstruct lost image information. Even when 90% of the data points were removed, the GAN network reconstructed the sample image with fewer artifacts, in contrary, the CNN model performed worse when most of the data were removed. Visually blurred reconstructed tissue and artifacts around the scanned object are noticeably in the presented case. By looking at the three most extreme corruption levels less edge preservation and structural details are visible at the results of the CNN model compared to GAN.

4. Discussion

In this paper, two image reconstruction approaches were compared for compensating lost data information in PET sinograms. The first approach was featured by a convolutional encoder-decoder network with four skip connections, providing the ability to reuse feature maps of first layers as the input to the deeper layers, preserving high-resolution features to the network, therefore, increasing the quality of sinogram reconstruction. In addition, the simple yet advantageous strategy was applied to use the pretrained network to train other networks; that is, the models for a higher level of corruption (30%, 50%, 70%, 90%) were trained by using weights from a model which was trained from scratch on 10% corrupted data. This appeared to be a very effective strategy since each model converged faster, without the necessity to adjust hyperparameters every new training. Hence, reusing weights allowed to achieve better reconstruction performance and reduce computational cost. The second approach was a recently proposed Generative Adversarial Networks. GAN had two networks; the discriminator contained four convolutional layers to classify on the 4x4 patches if the output sinogram was real or not and a generator which was previously used CNN to produce more and more realistic images. The advantage of GAN was that it learn a loss function that adapts to the data.
On the basis of the presented results, both CNN and GAN networks performed remarkably well, reaching high PSNR and SSIM scores, throughout both data representations on multiple levels of corruption. Such results can be explained by the beneficiaries provided by the raw sinogram format compared to post-reconstructed PET images. Sinograms are more uniform and provide a lighter data format consisting of angle based histograms, compared to edgy, noisy post-reconstructed PET images. Therefore, it is easier for the network to learn important features giving a promising performance. These results provide additional support to the Yu and Muhammed [11] findings where denoising of pre-reconstructed sinograms showed better results then denoising on post-reconstructed PET images.
The quantitative results (Figure 6) indicated that GAN scored higher CNN autoencoder with a small margin when predicting lost information based on all three metrics for both data representations. On average of all five levels of corruption, PSNR, and SSIM when predicting missing pixels were 41.44, 0.977 and 42.341, 0.983 for CNN, GAN, respectively. When predicting missing projections, CNN scored 40.13, 0.866 in contrary to GAN 46.84, 0.989. Even a small improvement in average PSNR and SSIM metrics indicated that the GAN was able to predict the removed information more accurately, than CNN. Additionally, graphical results indicated that GAN performance

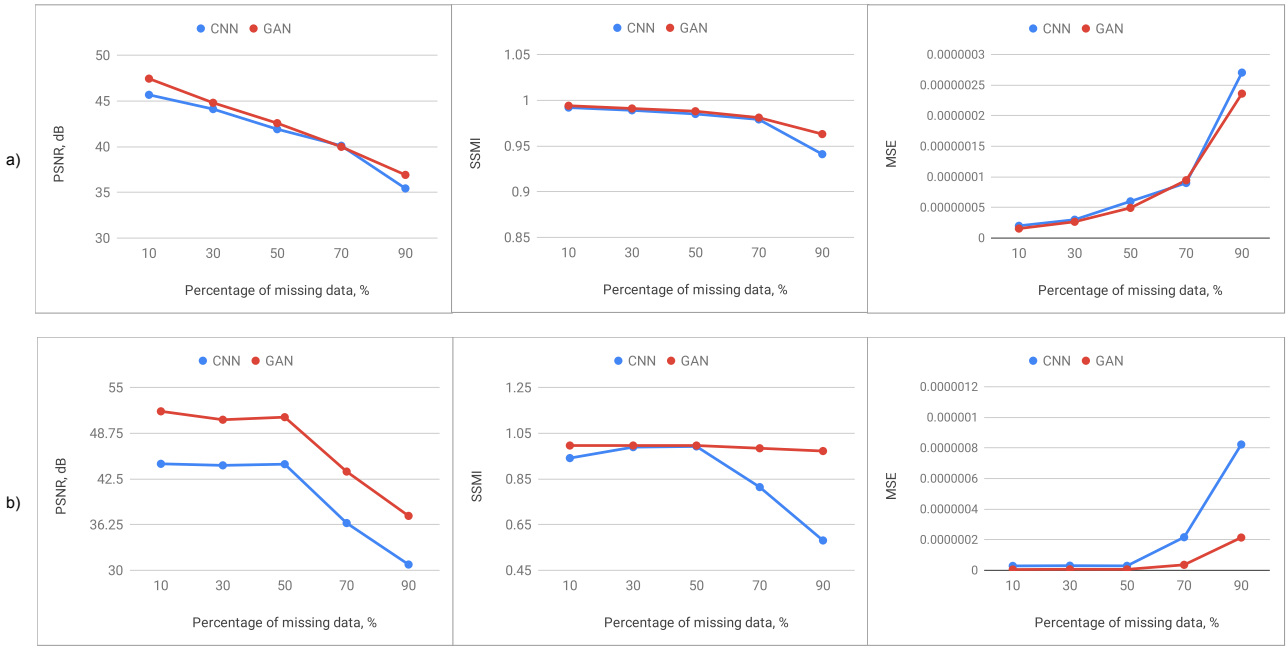

Figure 5: *Graphs of PSNR, SSIM and MSE scores dependency on the amount of missing data for two different data representations: (a) missing pixels, and (b) missing projections, using CNN and GAN approaches*

was more stable while CNN performance perceptibly decreased when more than half of that data were excluded (70,90%). CNN was using MSE as an objective function, which is a simple yet effective function when training autoencoder. However, multiple studies concluded that penalizing training with MSE tends to produce over-smoothed and blurred images, which eventually decreases metric scores [32, 29], while GAN objective function was more adaptive to the dataset.

The qualitative results confirmed that CNN blurred the reconstructed images at the corruption level of 70% and above, while GAN produced sharp and more realistic output even at the highest corruption level. Beside mentioned differences in the objective function of both models, additionally, GAN discriminator panelized structure at the size of the defined patch, while CNN did it on the full size of the image. According to original pix2pix paper [29], penalizing on the patch size is beneficial for GAN network as it makes training faster, requires fewer parameters and still outputs in high-quality results.

Strengths and limitations

The proposed reconstruction methods add up to the other studies which provide novel approaches of medical image reconstruction as an alternative for classical filters [33, 16]. The suggested framework has several strengths. First, this study used real PET dataset, which reflects real situations in the medical field: data contains complex noise, it is not consistent, therefore sometimes hardly readable. In contrary, the other studies used Shepp-Logan phantom to simulate sinograms for denoising and reconstruction problems [9, 11], which gives the advantage of controlling multiple simulation parameters. However, such data does not fully reflect the real-world scenario. Furthermore, this study suggested a novel end-to-end framework for using standard medical PET images in the way that the images were transformed into the sinogram domain, applied on the neural network model and converted back to the image space using filtered backprojection algorithm. Such a framework might be extended and used with other imaging modalities, such as CT, MRI, SPECT, when raw data is unavailable.

On the other hand, the proposed framework had some limitations too. When applying the filtered backprojection algorithm, there was visible noise induced on the PET image. However, the raw dataset in the sinogram domain was not available while the study was conducted; therefore, using Radon transform to

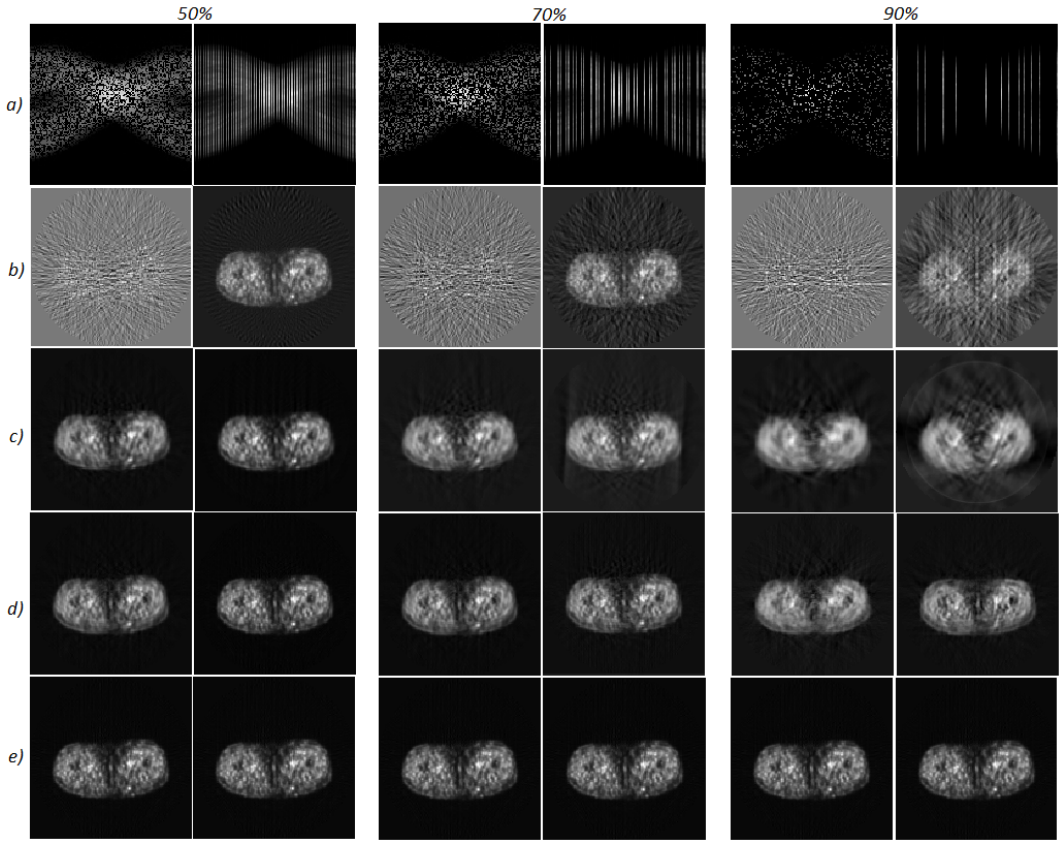

Figure 6: *Sample image from the testset in most extreme levels (50, 70, 90%) of corruption by removing projections and random pixels. (a) The corrupted sinogram, (b) FBP reconstruction without missing data compensation, (c) FBP reconstruction after compensating missing data with CNN, (d) FBP reconstruction after compensating missing data with GAN's, (e) ground truth image*

acquire sinogram dataset was the chosen solution. Consequently, this trade-off in decreased quality was inevitable, but it did not contribute to the study goals. Such limitation might be evaded be accessing raw data directly from PET scanner. Additionally, the more extensive hyperparameter tuning, using grid-search or random search techniques instead of manual tuning might have lead to achieving even more impressive results. However, the study had time constraints, and since there were many networks to be trained, more complex hyperparameter search was simple out of project scope.

5. Conclusion

In conclusion, based on the thorough search of the similar studies, this paper suggested the first attempt of applying the end-to-end framework to compensate simulated missing pixels and missing projections on PET sinogram domain, which is an existing problem of PET scanners, inducing the additional noise and artifacts. Two neural network based approaches were applied, compared, and evaluated quantitatively and qualitatively. More complex GAN model outperformed CNN autoencoder by scoring higher in both PSNR and SSIM metrics and promisingly performing even when 90% of data were removed. In the future, more experiments and testing should be performed to improve this initial work before translating such framework into the clinical setting.

Bibliography

[1] Feng Shi et al. " Prediction of standard-dose brain PET image by using MRI and low-dose brain [18 F]FDG PET images ". In: *Medical Physics* 42.9 (2015), pp. 5301–5309.

[2] Selena Milicevic Sephton and Simon M. Ametamey. *Positron emission tomography agents*. 2013, pp. 7–20.

[3] Hansen DC et al. "ScatterNet: a convolutional neural network for cone-beam CT intensity correction David". In: *Medical Physics* (2018).

[4] H.W.A.M. de Jong et al. "Correction Methods for Missing Data in Sinograms of the HRRT PET Scanner". In: *IEEE Transactions on Nuclear Science* 50.5 (2003), pp. 1452–1456.

[5] T. Leroy Laurence. *Dead pixel compensation in PET*. 2016.

[6] Biting Yu et al. "3D conditional generative adversarial networks for high-quality PET image estimation at low dose". In: *NeuroImage* 174.October 2017 (2018), pp. 550–562.

[7] F. Hashimoto et al. "Denoising of Dynamic Sinogram by Image Guided Filtering for Positron Emission Tomography". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 2.6 (2018), pp. 541–548.

[8] Kostadin Dabov et al. "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering". In: 16.8 (2007), pp. 2080–2095.

[9] Sari Peltonen et al. "PET sinogram denoising by block-matching and 3D filtering". In: *IEEE Nuclear Science Symposium Conference Record* (2012), pp. 3125–3129.

[10] M M Hasan. "Adaptive Edge-Guided Block-Matching and 3D Filtering (BM3D) Image Denoising Algorithm". In: May (2014), pp. 1–120.

[11] Sicong Yu and Hamed Hamid Muhammed. "Comparison of pre- and post-reconstruction denoising approaches in positron emission tomography". In: *Proceedings of 2016 1st International Conference on Biomedical Engineering: Empowering Biomedical Technology for Better Future, IBIOMED 2016*. 2017, pp. 1–6.

[12] Zepa Yang et al. "A deep learning-enabled iterative reconstruction of ultra-low-dose CT: use of synthetic sinogram-based noise simulation technique". In: 1057335.March 2018 (2018), p. 112.

[13] Le An et al. "Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI". In: *Neurocomputing* 267 (2017), pp. 406–416.

[14] Ian J. Goodfellow et al. "Generative Adversarial Networks". In: (2014), pp. 1–9.

[15] Ugur Demir and Gozde Unal. "Patch-Based Image Inpainting with Generative Adversarial Networks". In: (2018).

[16] Ji Zhao et al. "Unsupervised Learnable Sinogram Inpainting Network (SIN) for Limited Angle CT reconstruction". In: Dl (2018), pp. 1–10.

[17] Kenneth Clark et al. "The cancer imaging archive (TCIA): Maintaining and operating a public information repository". In: *Journal of Digital Imaging* 26.6 (2013), pp. 1045–1057.

[18] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. Vol. 13. 1. MIT Press, 2017, p. 35.

[19] M. Vallières et al. "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities". In: *Physics in Medicine and Biology* 60.14 (2015), pp. 5471–5496.

[20] Oleg S. Pianykh, Ksenia Pospelova, and Nick H. Kamboj. "Modeling Human Perception of Image Quality". In: *Journal of Digital Imaging* 31.6 (2018), pp. 768–775.

[21] J Johansson. "Radon transform in PET". In: 6.2 (2004), p. 103.

[22] Christopher M Bishop. *Pattern Recognition and Machine Learning*. Vol. 53. 9. 2013, pp. 1689–1699.

[23] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: (2014), pp. 1–15.

[24] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9351 (2015), pp. 234–241.

[25] Feng Shi et al. "Multi-Level Canonical Correlation Analysis for Standard-Dose PET Image Estimation". In: *IEEE Transactions on Image Processing* 25.7 (2016), pp. 3303–3315.

[26] Gao Huang et al. "Densely connected convolutional networks". In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-January (2017), pp. 2261–2269.

[27] Hang Zhao et al. "Loss Functions for Image Restoration With Neural Networks". In: *IEEE Transactions on Computational Imaging* 3.1 (2016), pp. 47–57.

[28] Hongming Shan et al. "3-D Convolutional Encoder-Decoder Network for Low-Dose CT via Transfer Learning From a 2-D Trained Network". In: *IEEE Transactions on Medical Imaging* 37.6 (2018), pp. 1522–1534.

[29] Phillip Isola et al. "Image-to-image translation with conditional adversarial networks". In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017* 2017-Janua (2017), pp. 5967–5976.

[30] Peter Aundal and John Aasted. "The Radon Transform Theory and Implementation Peter Toft Department of Mathematical Modelling Section for Digital Signal Processing Technical University of Denmark". In: (1996).

[31] Zhou Wang et al. "Image quality assessment: From error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612.

[32] Chuan Li and Michael Wand. "Precomputed real-time texture synthesis with markovian generative adversarial networks". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9907 LNCS (2016), pp. 702–716.

[33] Yan Wang et al. "Predicting standard-dose PET image from low-dose PET and multimodal MR images using mapping-based sparse representation". In: *Physics in Medicine and Biology* 61.2 (2016), pp. 791–812.

PET Image Reconstruction using Convolutional Neural Network and Generative Adversarial Network in Sinogram Domain

-Worksheet-

10th semester 2019

Biomedical Engineering & Informatics
Project group: 19gr10418

Ignas Kupcikevičius

Aalborg University
School of Medicine and Health
Fredrik Bajers Vej 7
DK-9220 Aalborg

AALBORG UNIVERSITET
STUDENTERRAPPORT

Title

PET Image Reconstruction using Convolutional Neural Network a nd Generative Adversarial Network in Sinogram Domain

Project

P8 - Biomedical (information) systems

Project period

10th semester

Project group

19gr10418

Participants

Ignas Kupcikevicius

Supervisors

Main supervisor:
Alex S. Jørgensen

Co-supervisor:
Anders Bertil Rodell

Pages: (PDF: 36)
Number of appendixes:
Submitted: 07/06/2019

Abstract:

Background: Artifacts and noise in PET imaging are caused by multiple factors, including lowered radiotracer dose and information loss in the form of missing pixels or missing projections. Recently, deep learning based algorithms have achieved promising results in the medical imaging field, including PET denoising or CT sinogram inpainting, especially using the Convolutional Neural Network (CNN) and Generative Adversarial Networks (GAN) architectures. This article aims to compare CNN and GAN approaches for PET sinogram missing data reconstruction task.
Methods: The end-to-end framework, from PET image to the sinogram domain and back to PET image domain, was proposed. The Radon transform was applied to covert PET images into sinograms. The first model was the CNN encoder-decoder based network with four skip connections. The effective strategy was applied to efficiently train more corrupted PET sinograms by loading previously trained weights. The second approach was the GAN network, with the generator designed similarly as the CNN encoder-decoder, and the discriminator containing four convolutional layers to classify generated sinograms as artificially generated or ground truth. The proposed framework ended by applying filter back projection algorithm to transform sinograms back to PET image domain.
Results: The results revealed that GAN outperformed CNN by a small margin. The average PSNR and SSIM scores within all five corruption levels were 41.44, 0.977, and 42.34, 0.983 when predicting missing pixels. Differences of two metrics between CNN and GAN were higher when predicting missing projections; 40.13, 0.866 versus 46.84, 0.989. Additionally, GAN performed noticeably better when 90% of sinogram data were removed, resulting in a sharper and more detailed reconstructed image, qualitatively comparing to CNN.
Conclusion: Different network architectures, chosen, and objective functions, might be the reasons why GAN performed better than CNN. Even though the study had some limitations, the promising results were achieved, which motivates to experiment further.

Acknowledgements

Appreciation is given to the supervisors Alex S. Jørgensen, and Anders B. Rodell for their contributions and guidance throughout the project.

Contents

SUMMARY

Noise in positron emission tomography (PET) has attracted major attention in the medical imaging field since it directly affects the diagnostic capabilities of PET modality. Artifacts and noise in PET imaging are caused by multiple factors, including lowered radiotracer dose and information loss in the form of missing pixels or missing projections. Recently, deep learning based algorithms have achieved promising results in the medical imaging field, including PET denoising or CT sinogram inpainting, especially using the Convolutional Neural Network (CNN) and Generative Adversarial Networks (GAN) architectures.

This article aims to compare CNN and GAN approaches for PET sinogram missing data reconstruction task. For that purpose, two existing PET problems were simulated, including missing pixels and missing projections of five levels of corruption.

The end-to-end framework, from PET image to the sinogram domain and back to PET image domain, was proposed. The Radon transform was applied to convert PET images into sinograms. The first model was the CNN encoder-decoder based network with four skip connections. The effective strategy was applied to efficiently train more corrupted PET sinograms by loading previously trained weights. The second approach was the GAN network, with the generator designed similarly as the CNN encoder-decoder, and the discriminator containing four convolutional layers to classify generated sinograms as artificially generated or ground truth. The proposed framework ended by applying filter back projection algorithm to transform sinograms back to PET image domain.

The results revealed that GAN outperformed CNN by a small margin. An average PSNR and SSIM scores within all five corruption level were 41.44, 0.977, and 42.34, 0.983 when predicting missing pixels. Difference between two metrics using CNN and GAN were higher when predicting missing projections; 40.13, 0.866 versus 46.84, 0.989. Additionally, GAN performed noticeably better when 90% of sinogram data were removed, resulting in a sharper and more detailed reconstructed image, qualitatively comparing to CNN.

Different network architectures, chosen, and objective functions, might be the main reasons why GAN performed better than CNN. Even though the study had some limitations, the promising results were achieved, which motivates to experiment further. This study showed how CNN and GAN performance degraded when more and more data were excluded. It also revealed how effective Generative Adversarial Networks works in terms of image reconstruction task. However, the possibles and implementation of such algorithm remained unexplored and left for future research.

Chapter 1

Background

In this chapter, the positron emission tomography imaging and related problems are presented. The existing and possible solutions are discussed, including machine learning based approaches.

1.1 Positron Emission Tomography Scans

Positron Emission Tomography (PET) is one of the major imaging modalities, broadly used in hospitals for diagnostic and intervention purposes [1]. As a nuclear medicine, functional imaging technique PET reflects the metabolism changes by using the injected radioactive tracers in the human body. Different radiotracers might be used to utilize PET in oncology, cardiology, neuroimaging, and other cases. Particularly, the PET system detects pairs of gamma rays emitted from the injected radioactive tracers on the biologically active molecule. The tracer concentration within the human body is recorded and 3D volume is generated using computer analysis [2].

1.2 Sinogram formation in PET

The scanner computing unit processes the raw data which are recorded during PET acquisition and stored in the sinogram format. There are two common ways to collect PET raw data, which are classified as 2-dimensional (2D) and 3-dimensional (3D) PET. Some scanners can gather the data in both 2D and 3D settings, whereas, some of them acquire data only in 3D or 2D. [3] Since the data used in this project are 2D PET images, the 2-dimensional data gathering type will be described further.
Firstly, the PET radiotracer that localizes the lesion in the patient body is administrated. Radiotracer allows emitting the protons from the tumor during the acquisition. Emitted protons instantly find an electron; this causes annihilation. When such protons annihilate with the electrons, the mass-energy is produced and converted into two gamma rays. Two gamma-ray photons with momentum are traveling in opposite directions and are registered by the PET detectors. The process when both annihilation photons reach detectors without interacting with surrounding atoms is called the coincidence event [2]. An example of the coincident event is visible in (Figure 1.1)

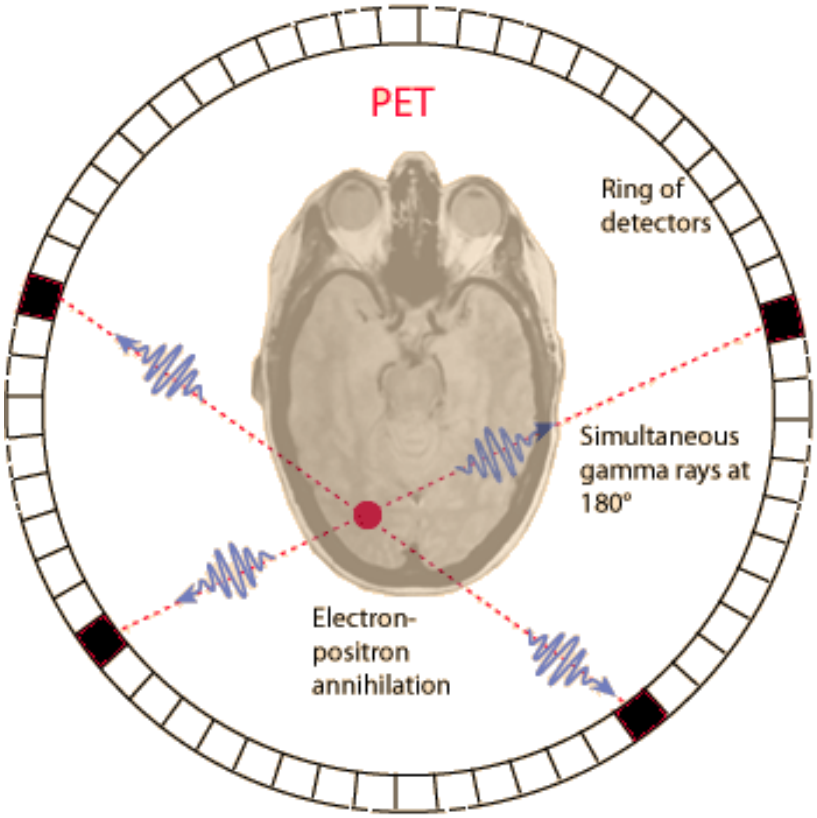

Figure 1.1: True coincidence event (Red dot) appears then positron-electron annihilation is recorded by a pair of detectors. Edited from [**?**]

A line between two detectors categorizes the coincident event. According to the literature, such a line is called a line of respond (LOR) [2]. By plotting multiple LORs as the functions of their angular orientation versus their displacement, the sinogram is formed. The displacement is the shortest distance between the LOR and center point (marked as X in Figure 1.2). As shown in Figure 1.2a) four coincidence even detections are characterized by 4 LORs (Figure 1.2 A,B,C,D) at different angles, and distances to the center. When multiple numbers of such LORs are plotted from the same point, the sine wave is formed (Figure 1.2 b)). A large number of coincidence events forms a sinogram (Figure 1.2 c))

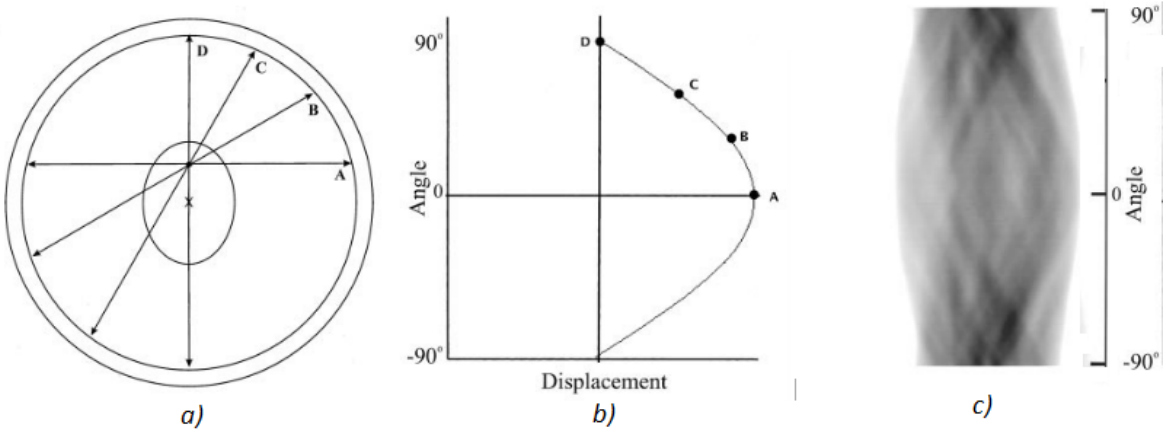

Figure 1.2: True coincidence event (Red dot) appears then positron-electron annihilation is recorded by a pair of detectors. Edited from [3]

PET scanner directly acquires the data into sinogram format, when each detector pair corresponds to a particular pixel in the sinogram, based on its orientation angle and the distance to the center. Thus, for every coincidence event detection, the LOR is determined, and the pixel associated with that LOR is projected into the final sinogram. Each 2D PET slice is separate sinogram acquired using a similar principle. [3].
In sinogram, the values along the horizontal rows represent the LORs at the particular angle, and are called the "projections". Every row in the sinogram is a projection through the object at the specific angle.[3]

1.3 Existing limitations and solutions in PET

PET images usually contain more noise and artifacts compared to magnetic resonance imaging (MRI) and computed tomography (CT). The low number of coincident events imposed by a reduced dose of a radiotracer and faulty electronics which causes the loss of information induces additional noise and artifact in PET.
High-quality PET images are often used in diagnosing brain diseases and disorders, such as Alzheimer's, dementia [2]. However, to capture high-quality PET images, usually, a normal dose of radioactive tracer needs to be used, which raises the concerns about possible health hazards [4]. In the report "Biological Effects of Ionizing Radiation (BEIR VII)" the estimated risk increment is about 0.04% per single PET scan, which multiplies after repetitive examinations. On the other hand, if the radiation dose is reduced, the quality of resulting images is degraded, due to the lower amount of detected photons Shi et al. [1], which negatively affects the diagnostic capabilities of PET. The difference between low-dose and the normal dose is presented in Figure 1.3 Thus, the necessity of the method to denoise low-dose PET image in the way the resulting image would be similar to the normal-dose PET image quality arises. Additionally, a comparison study by Yu and Muhammed [5] suggested that signal-noise ratio (SNR) was improved by sinogram-based PET reconstruction.

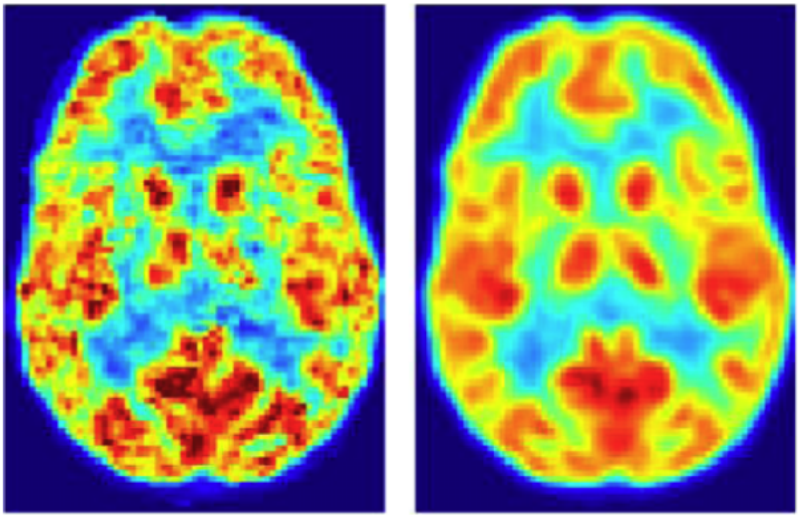

Figure 1.3: Example image of low-dose PET image versus normal-dose PET image. Edited from [2]

Recently, PET systems started to use silicon photomultipliers (SiPMs) as a detector for photon registration [6]. According to the literature, there is a high probability that some of the SiPMSs may not function properly, resulting in dead pixels appearance. A dead pixel is a pixel with a missing data value, which causes the degraded quality of reconstructed PET images (Figure 1.4 b)). Additionally, a dead pixel might arise due to the problems in processing electronics [6]. The study by de Jong et al. [7] investigated a few methods to compensate for missing data caused by the gaps between high-resolution PET scanner detectors and faulty parts (Figure 1.4 a). Bilinear and the model based interpolations were applied and compared, where model-based approach outperformed bilinear interpolation, but with a price of long processing time. [7]

A variety of studies used commonly know algorithms for improving PET image quality. Firstly, the Gaussian filtering and bilateral filtering were used to improve PET post-reconstruction; however, such filtering resulted in reduced edges preservation and blurred images [8]. Furthermore, the block-matching 3D (BM3D)[9], which is a variation of non-local mean denoising approach [10] was applied to PET images, showed structural improvements [11]. However, the study used simulated, not the real PET data, while BM3D is not effective on natural images due to the lack of highly correlated matching blocks.[12]. Another attempt to improve medical imaging quality is to apply a filter on the sinogram domain instead of reconstructed images. Sinogram is an angle based histograms which is the raw data format in PET systems.[2].

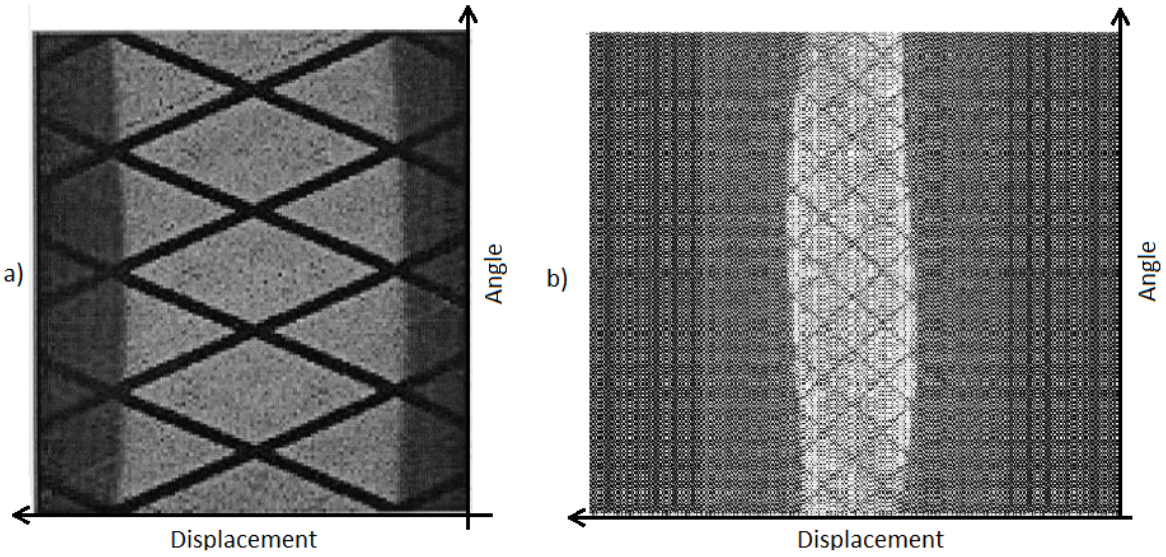

Figure 1.4: Example image of corrupted sinograms. Edited from [7, 6]

1.4 Common Techniques for PET Image Quality Improvement

Multiple methods were proposed in order to reduce the noise in PET images and to improve the overall image quality while maintaining image details. The study by Bagci and Mollura [13] uses the singular value thresholding concept and Stein's unbiased risk estimate to optimize a soft thresholding rule. Another study by Visvikis et al. [14] considered to combine the complementary wavelet and the curvelet transforms to address the issue of resolution loss which usually appears with standard denoising techniques. Finally, Ortega Maynez et al. [15] proposed a multi-resolution method for PET image denoising in the transform domain by modeling each sub-band separately as a group of different regions separated by boundaries. Besides mentioned techniques which mainly focused on normal-dose PET image denoising, there are few recent studies which considered to reconstruct low-dose PET image. Such an approach is useful, due to the public concerns about the health hazards of radioactive tracers. One such study was conducted by Shi et al. [1], who trained a random forest to estimate a full-dose PET image from low-dose. The study concluded that standard-dose PET can be well predicted from low-dose scans and MRI combination. Similar conclusions were generated by author Wang et al. [16], who used a mapping based sparse representation to reconstruct full-dose PET images by utilizing both low-dose PET and multimodal MRI. The same year, Shi et al. [17] proposed a multi-level canonical correlation analysis framework to connect the low-dose and normal-dose PET images into a common space and then performed patch-based sparse representation for the estimation, which showed the superior performance compared to other methods. However, there were some limitations too. One of them is that such spare learning based method was performed on small patches which makes

the voxel-wise estimate strategy very time-consuming. Moreover, the over-smoothing effect was noticed on qualitative results due to the averaging of the overlapped patches.

1.5 Neural Network Based Techniques for PET Image Quality Improvement

Even though the common techniques improve PET quality, they run slowly, impose additional noise. Most of the time using such algorithms, require human interaction to define features manually. An idea to automatically learn important features, which would help to reconstruct missing data or to remove noise is possible by using state-of-the-art neural networks approach.

1.5.1 Convolutional Neural Network

Convolutions Neural Networks (CNN) has become a state-of-the-art technique in terms of image analysis. It perfectly fits imaging data due to the convolving layers which help to reduce dimensions of the input data. This CNN feature is useful for most of the natural images due to their high-resolution pixel space [18]. Medical imaging is a field where image resolution grows rapidly due to the technological improvements and the health-related importance of high-quality image capabilities.
The study by An et al. [19] proposed a deep auto-context CNN for full-dose PET image estimation based on the local patches from the low-dose PET and T1 weighted MRI. Worth to mention, that the author conducted the study by extracting the 2D slices from the 3D volumes, resulting in the lost spacial information and limiting the generalization of such technique.

1.5.2 Generative Adversarial Network

Generative Adversarial Networks (GAN) have recently received widespread attention [20, 21, 22]. GANs know as generative models which have two main units: generator and discriminator. The generator learns to replicate the input images, while the discriminator tries to distinguish between the generated and real datasets. Using such an approach, many researchers achieved promising results. [21, 22]
In terms of PET data, the most recent study which used GANs in a 3D setting was conducted by Yu et al. [4]. The study proposed a conditional GANs to estimate the high-quality full-dose PET images from low-dose ones. Specifically for the generator, they used the famous 3D U-net-like deep architecture, which can combine hierarchical feature by using skip connection.

Chapter 2

Methods

In this chapter, the data used in this study, and the methods applied to reach the study goals will be presented.

2.1 Study design

The study workflow consists of 5 different steps: *data acquisition, pre-processing, neural network generation, post-processing, model evaluation, and validation.* By describing each of them, the transparency of the study shall be maintained and the necessary information provided to the reader.

2.2 Data acquisition

Since there were no specifically collected data for this project, the publicly available dataset was acquired from the Cancer Imaging Archive (TCIA: http://www.cancerimagingarchive.net), an archive of the medical images of cancer created by the National Cancer Institute (NCI) [23]. Soft-tissue-Sarcoma dataset was chosen from TCIA, which is, based on the thorough search of the relevant datasets yielded a large number of tested subjects, giving the high number of training samples, necessary to train a deep neural network [24]. The dataset consists of 51 patients with histologically proven soft-tissue-sarcomas of the lower extremities in

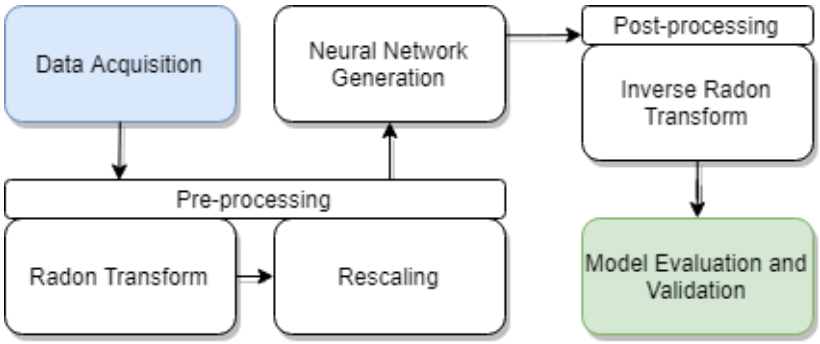

Figure 2.1: Proposed study workflow which starts by gathering data (blue color) and ends by evaluating model performance (green color).

PET/CT/MRI image formats [23, 25].

For this study PET data from all 51 subjects were used which contained DICOM (Digital Imaging and Communications in Medicine) images of 128 x 128-pixel resolution (Figure 2.2) in a total of 13,417 2D images.

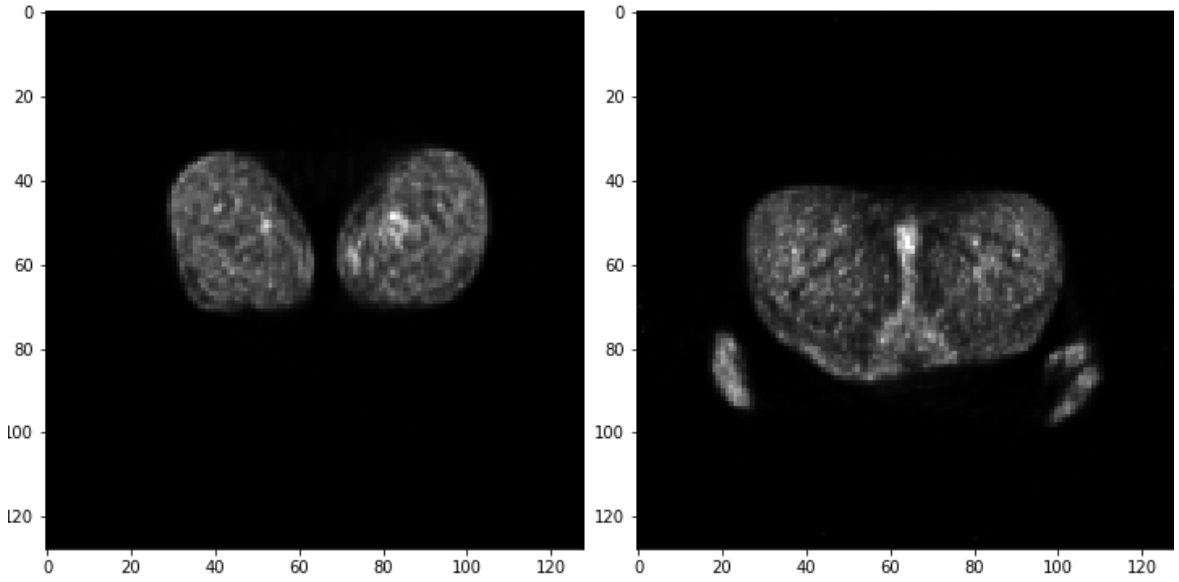

Figure 2.2: PET images of Soft-tissue-Sarcoma of lower extremities .

The fludeoxyglucose (FDG) radiotracer was used when performing PET scans on a PET/CT scanner (Discovery ST, by GE Healthcare). A median of 420 MBq of FDG was injected intravenously followed by the 60 min of body imaging acquisition.[26] Additionally, the attenuation correction was performed [26] which reduces severe artifacts induced due to the high number of lost photons during the procedure [27].

2.3 Radon transform

Since the data from TCIA was obtained in standard imaging format, meaning it was already reconstructed, data needs to be transformed back into the raw format to work in the sinogram domain. This might be achieved by using the Radon transform. Radon transform of the image described as the function $f(x, y)$, is defined by the combination of the integrals through $f(x, y)$ in vertical and horizontal projections, or y and x axis, respectively.[28] Radon transform $R(p, \tau)$ (Figure 2.3) mathematical is defined as:

$$Rf = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(xcos\theta + ysin\theta - s)dxdy$$

where θ is the angle of the line and s is a perpendicular offset of the line.[28]

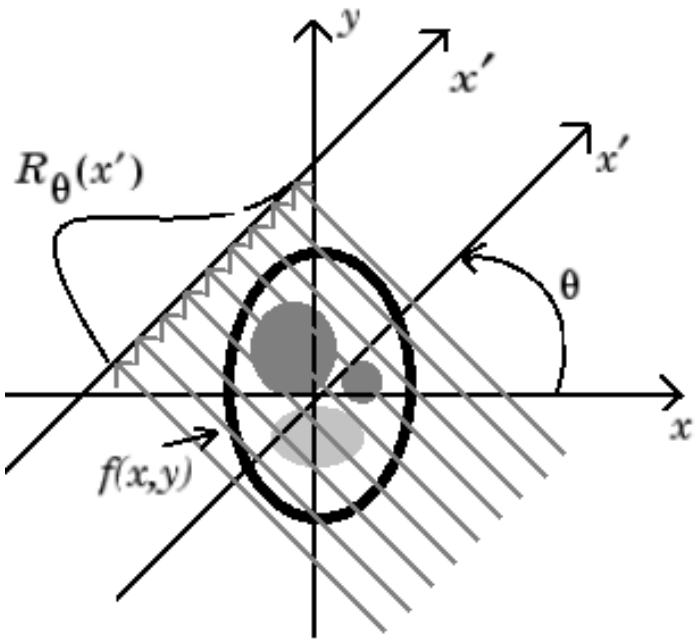

Figure 2.3: Fundomentals of Radon transform. Edited from [29]

In this study the data was transform by using *skimage* library in Python coding environment. *Skimage* allows to define two parameters before Radon transformation: the angle *theta*, which was left as a default value of 180, and *Circle*, which when set on 'True' allows to make a uniform sinogram size of 128 projections and 128 detectors, corresponding to the size as the original image. This property is helpful since CNN requires uniform dataset (Figure 2.4).

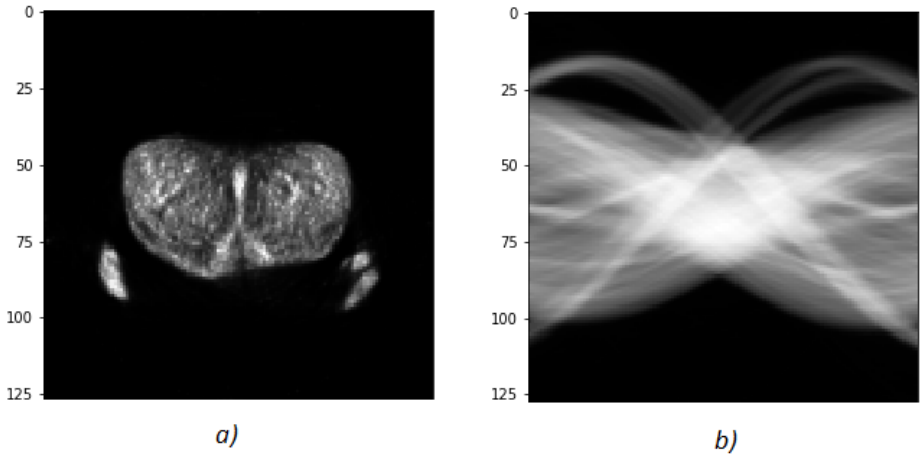

Figure 2.4: Radon transform applied on the PET image (a). Resulting sinogram (b) is used for further pre-processing.

Two data representations were created to address existing problems in PET imaging domain: dead pixels and missing lines of response (LOR). Dead pixels are simulated by removing random data points from the PET sinogram. The amount of randomness is increased in five levels, 10%, 30%, 50%, 70% and 90% (Figure 2.5 a)), to reflect multiple scenarios of missing patterns, also, to investigate how far CNN and GAN can perform to properly compensate missing pixel values. The second data representation is focused to simulate missing projections, which appears in PET sinograms due to the gaps between detector heads in the scanner or because of the faulty detectors [7]. Also presented in five different levels (Figure 2.5 b))

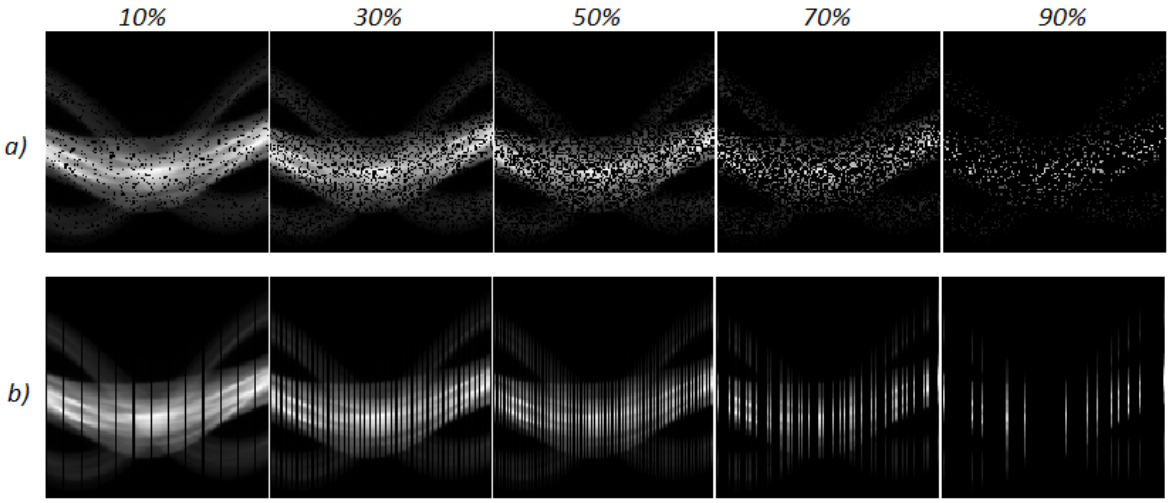

Figure 2.5: Representations of corrupted sinograms with simulated dead pixels a), and missing projections b) at three different levels.

2.4 Intensity normalization

Applying the Radon transform converted the whole data into the sinogram domain. Therefore, further pre-processing was proceeded to prepare data for the neural networks. This involved data rescaling and reshaping.
Data rescaling helps to control the weights of the data not to become large values during the training phase. The large weight values usually result in the poor performance and unstable model, thus, it is critical to rescale it before presenting data to the network. [30] According to the literature, one of the most often used rescaling methods is to make the data elements lie in the range between 0 and 1, [30] thereby, this was achieved by applying the following equation:

$$Z_i = \frac{Y_i - Y_{min}}{Y_{max} - Y_{min}}$$

where Z_i is rescaled data, Y_i original data [31].

Even though convolutional neural networks have become very popular between the researchers, it came with a few restrictions. One of them is the fixed size of that input data Ghosh et al. [32]. Different deep learning libraries in Python requires a different shape of the input tensor. For example, PyTorch [33] library expects that the first element of the input data would correspond to the number of channels, while in Keras [34] it should stand last. Since this study was using Keras library, data was reshaped it the way they the first, second and third dimensions corresponds to X, Y, Number of Channels, respectively.

2.5 Model generation

In this paper, two different architectures are presented and compared for missing data reconstruction in PET sinogram domain. The first architecture is a convolutional neural network based autoencoder (Figure 2.6) inspired by U-net CNN for image segmentation [35]. Proposed network has 4 convolutional layers with 32, 48, 64, 64, respectively. These layers encode the information into a dimensionally smaller data form, containing automatically learned most important features. Then, the compressed data is fed through 4 deconvolving layers of 64, 64, 48, 32 neurons, respectively. Deconvolutional layers of the network might be called a decoder part. It tries to reconstruct the compressed image data into the normal size (128 x 128) image. 3 x 3 filter size was used through the entire network, with a stride of one, based on the similar studies [35, 17]. Additionally, 4 skip connections were added which shuttle low-level features to the high level features [22]. Using skip connections noticeably increased reconstructed image structural details. Finally, based on previous experience and other studies Shan et al. [36], it was decided not to use max-pooling layers, which originally exists in U-net architecture. Using max-pooling layers results in extra image distortions which appear after reconstruction. Similar artifacts were found in mentioned in a study by.Shan et al. [36]

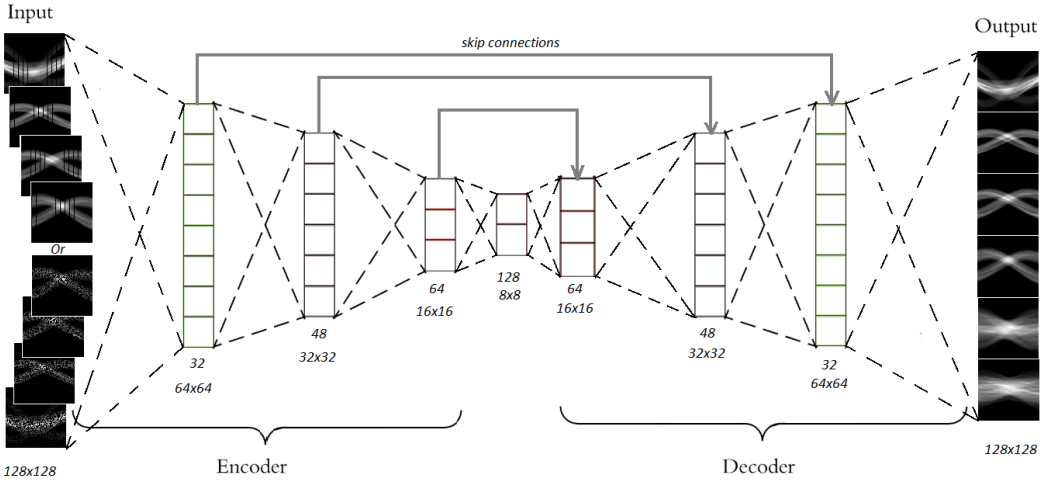

Figure 2.6: Autoencoder CNN architecture. Input layer contains corrupted sinograms, output layer fully reconstructed sinograms.

The second architecture is based on recently proposed Generative Adversarial Networks (GANs) [21, 37, 22]. Unlike the common neural networks, in GANs consist of two models: generator and discriminator. The GAN automatically learns the goal specific lost function, which classifies the output into real (ground truth) or fake (artificially generated), while at the same time trains the generative model to minimize this error Isola et al. [21]. Blurred images will be ranked as 'fake', therefore, the GAN will try to adjust generative model weights further in order to generate more realistic images. This is the biggest advantage against normal CNN and its pre-defined lost functions, such as MSE, where network tries to minimize the loss and ends up with blurred, unrealistic results. In this paper, the proposed GAN is based on the 'pix2pix' network (Figure **??**) [21], since it was previously successfully applied for the image inpainting problem and showed promising results [37]. Here, a generator is similar as described autoencoder CNN with skip connections, while the discriminator consists of 4 convolutional layers with 64, 128, 256, 512 neurons, respectively. The final output layers are also CNN with one neuron only followed by Sigmoid function, since the results are in binary form - label for real or fake image.

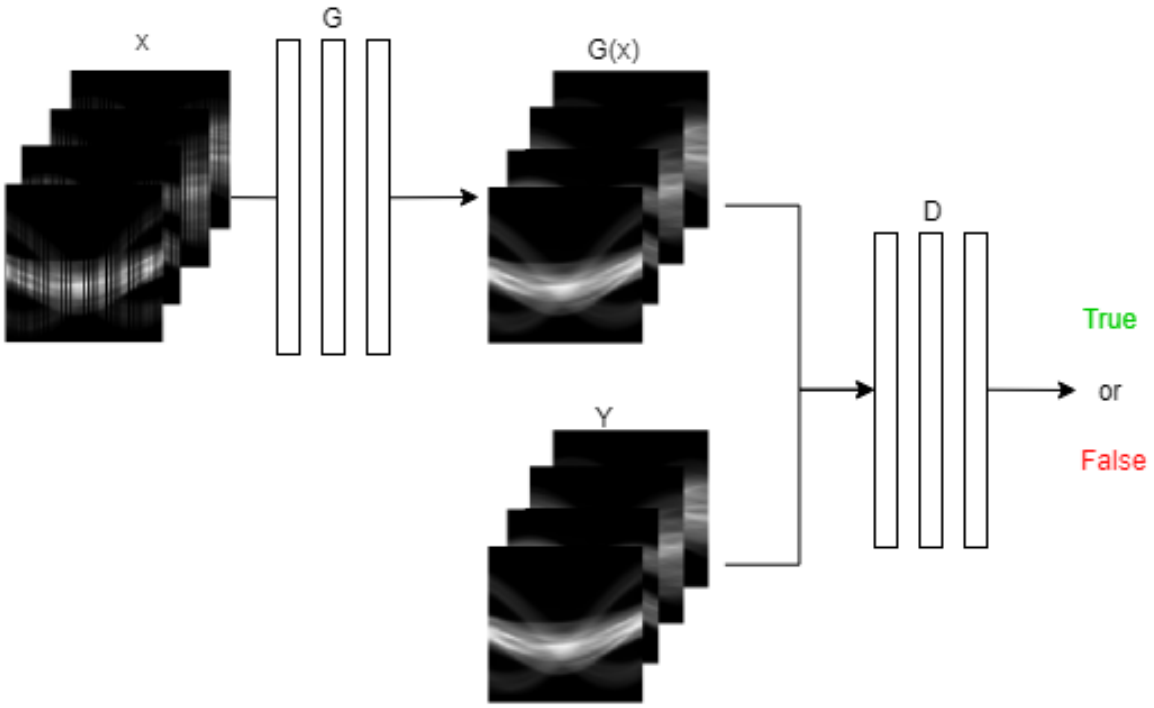

Figure 2.7: Training a GAN by passing corrupted images (x) through the generator (G) to synthesize indistinguishable images (G_x) from the real images (Y). The discriminator (D) learns to classify between real and fake images, while generator tries to fool discriminator

Most of the hyperparameter tuning was performed manually, where some of them were chosen according to the literature. Since the project aims to compare two different neural

network architectures with multiple data representations, using more advanced tuning techniques, such as grid search, or random search would simply take too much time, therefore, it was out of the project scope. Rectified linear unit (ReLu) as an activation function was used throughout the proposed network since it mathematically simple, yet effective, and recommended according to the literature [24]. Different optimizers were tested, however, Adam solver [38], with learning rate of 0,003 and, momentum parameters $\beta 1 = 0.5$, $\beta 2 = 0.999$, showed the fastest convergence and best results.
Making a comparison study requires to train multiple neural networks with different data representations, which is the overwhelming task itself due to the time it consumes, moreover, the amount of adjustable hyperparameters grows with each separate training. To overcome this problem, the transfer learning-based strategy was used which initializes the new model from pretrained one. [36] In our case, it would be enough to train from scratch one model with data representation where, for example, 10% projections, are missing. After desirable results are reached, the training weights are saved and might be loaded when training on more corrupted data with 30% and 50% of missing data. A similar plan is applied with second data representation where random pixels are removed instead of the projections. In this way, for both CNN and GAN's, only four models need to be trained from scratch compared to 12 different in total, which are compared in this study. A study by Shan et al. [36] have used this transfer learning strategy and by experimenting confirmed that performance through training based on transfer learning compared to the training from scratch showed better results and faster convergence.

Objective functions

The training and validation loss after every batch is computed between the batch of predicted and the ground truth values using Mean-square-error (MSE) also known as L2 loss, for CNN model. MSE is one of the most used objective function and it is available by default on Keras library. It was used in the CNN model since it is a *de facto* standard objective function in neural networks [39].

For GANs model the objective function was left as it was defined originally in 'pix2pix' paper [21], and mathematically described as a combination of GAN objective and L1 loss:

$$G^* = arg \min_G \max_D \mathcal{L}_{GAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

where G^* is a final objective, G a generator, D a discriminator, L_{GAN} a Laplace transform of GAN objective, L_{L1} a Laplace transform of L1 loss. G tries to minimize the objective function against the adversarial D that acts opposite and maximizes it [21].

2.6 Inverse Radon Transform

One of the easiest ways to reconstruct the image from the sinogram domain is by using filtered backprojection (FBP). When applying backprojection the blurring effect is induced

on image space. To overcome this limitation the filtering is used, which corrects blurring effect to some extent. [40] In this study different available filters were used and tested to see which one performs better for TCIA dataset. The ramp filter showed noticeably better results, hence, it was used when transforming the predicted test date data back to the image space. Even when filtering applied, there is visible noise induced after backprojection from the sinogram into image space (Figure 2.8).

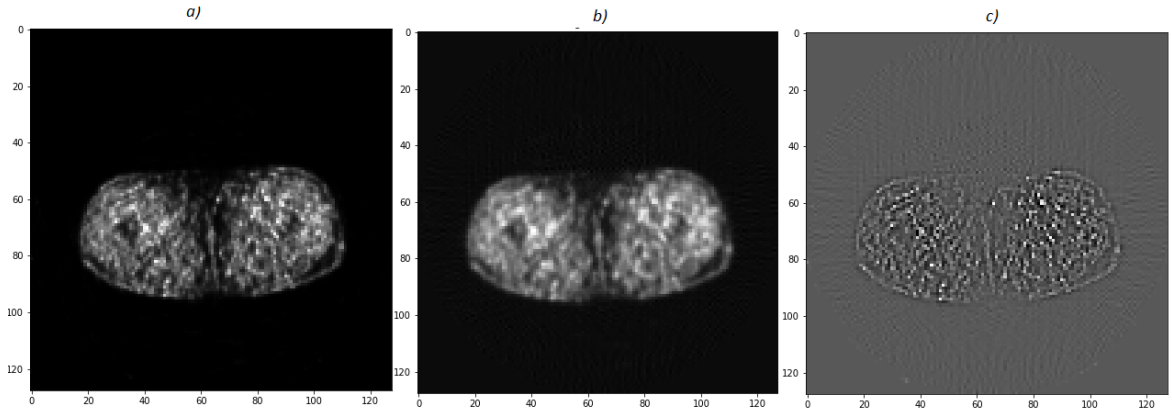

Figure 2.8: The lost image details due to Radon and Inverse Radon transforms, where orignal image a), FBP reconstructed image b), and the lost details c)

The raw dataset in sinogram domain was not available while the study was conducted and using Radon transform to acquire sinogram dataset was the only solution, therefore, this trade-off in decreased quality was inevitable, but it did not contribute to the study goals.

2.7 Evaluation and validation

In this section, used quantitative metrics are presented, followed by the strategy of how the model was evaluated.
From 51 data subjects, 5 random patients were put aside and left as the unseen test dataset, which approximately corresponds to the 10% of the total dataset. This test set is used only for the final results. The following 46 subjects correspond to the 12,082 sinograms were divided into training and validation datasets with a ratio of 70% to 30%, respectively. Training and validation data sets were used during the training phase and for hyperparameter tuning. After each training phase, the qualitative results can be observed for evaluation of reconstructed images, in terms of visible structural details, artifacts, blur. Qualitative results are important in the real-world scenario since this is how the doctors perceived image information in the working environment. Any image distortions increase the risk of misdiagnoses, thus, the quality of reconstructed images have a direct impact on the patient treatment [41]. Observing image quality visually might be very subjective evaluation, therefore, the quantitative metric should be introduced. Commonly used metrics within different imaging studies are peak signal-to-noise ratio (PSNR) and signal-to-noise ratio (SNR).[17, 36] PSNR is the

ratio measure between the power of image peak value and the power of the corruption that affects its representation, measured in decibels (dB) [42]. PSNR can be defined using the mean square error (MSE) measured between ground truth and corrupted representations:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2$$

where m, n stands for image dimensions, I for monochrome image and K for its corrupted approximation, i, j coordination system values. Knowing that the signals might have a wide dynamic range, PSNR could be expressed in the logarithmic decibels (in dB):

$$PSNR = 10 \cdot log_{10}(\frac{DR^2}{MSE})$$

where DR is a dynamic range of the image. Usually, the greater value of PSNR means a higher quality reconstruction of the image. [17]
Signal-to-Noise ratio is a ratio of the power of a signal to the power of noise and can be defined as follows:

$$SNR = \frac{P_s}{|x[n] - s[n]|^2}$$

where P_s is a discrete power of signal, and $|x[n] - s[n]|^2$ is a squared error between the original and corrupted images. Similarly as in PSNR, here the higher SNR value the better quality of the reconstruction. [17]

However, neither of PSNR or SNR do guarantee the best perceptual and textural outcome, even the resulting score is high. Both metrics measured per-pixel intensities, which is not a way of how a human perceives the image information. Additionally, according to the literature, these metrics suffer from the regression-to-mean problem resulting in blurred reconstructed images. Thus, an additional metric such as structural similarity measure (SSIM) Wang et al. [43] is also used.
SSIM considers image corruption level as perceived change in image structural information measures between two images. SSMI compares groups of pixels between two images, using three composite measure: luminance, contrast and structural. Simplified SSMI mathematical expression can be seen as:

$$SSIM(A, B) = l(A, B) \cdot c(A, B) \cdot s(A, B)$$

where A, B are groups of pixels from two images, l - luminance measure, c - contrast measure and s - structural measure. Since it compares pixel patches, not individual pixels like PSNR and SNR, SSIM is a metric which represents results more similarly as we observe. [43] Qualitative and quantitative results are evaluated using an only test set of 1335 reconstructed images.

Chapter 3

Results

In the results chapter, three different sections are presented. The first section is the validation results; the second section covers the quantitative results of reconstruction approaches. After that, the third section is for the qualitative findings.

3.1 Validation results

This comparison study was conducted by training 10 networks in total for both data representations and all five levels of corruption. In this section the convergence graphs are presented from the 10% only, to present how model performance was evaluated.
Both CNN and GAN were trained and validated using similar datasets. The training set contained 8 457 randomized sinograms and validation contains 3 625 randomized sinograms. After each training procedure, the model convergence graphs were printed to observed if model overfitted or underfitted. Additionally, to decide how many epochs are enough for the model to generalize well.
In addition, during the training, the defined batches of 32 sinograms were fed into the network and shuffled every epoch to ensure that highly correlated data would not be obtained in each batch. According to the author Ian Goodfellow et al. [24], shuffling the data like this results in fewer oscillations during the training. The neural network is not learning the specific order of the data, therefore, it converges faster. Below (Figure 3.1), the convergence graphs indicate training and validation loss performance during the training phase. The objective function used for training was Mean-square-error.

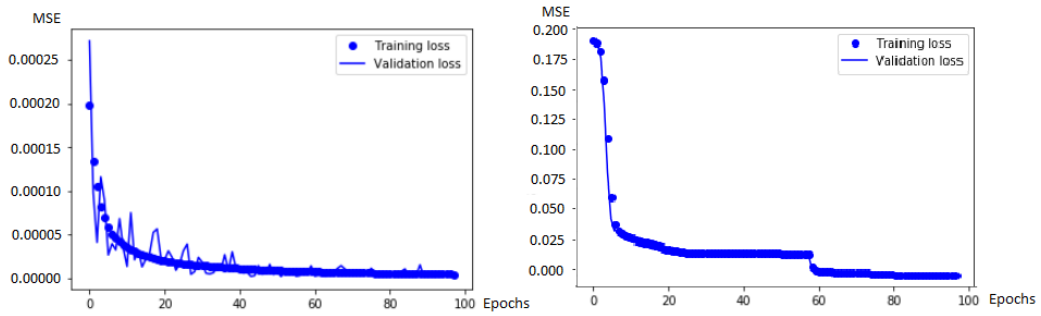

Figure 3.1: Convergence graphs of CNN model losses of training and validation datasets. Left graph represents training on missing pixels, right - training on missing projetions.

From the graphs, it is visible that the CNN model had a good performance on both data representations. There are no indications of overfitting or underfitting since the gap between the training and validation plots is relatively small. During multiple runs of manual hyper-parameter tuning, it was noticed that after 100 epochs training and validation loss does not decrease noticeably and are the MSE is a relatively small number. Therefore, the number of 100 epochs was selected for the final training.

The performance of the GAN is visible in (Figure 3.2). It is important to note that GAN during the training does not use validation dataset. It penalizes the predictions using the discriminator model to adjust the generator weights accordingly. From both graphs, it is visible that generator loss was decreasing, which means that the generator was able to produce more realistic images. Discriminator loss on the real images is fluctuating around 0.5 value, which means that discriminator guesses if the image is real or generated, which is a good indication.

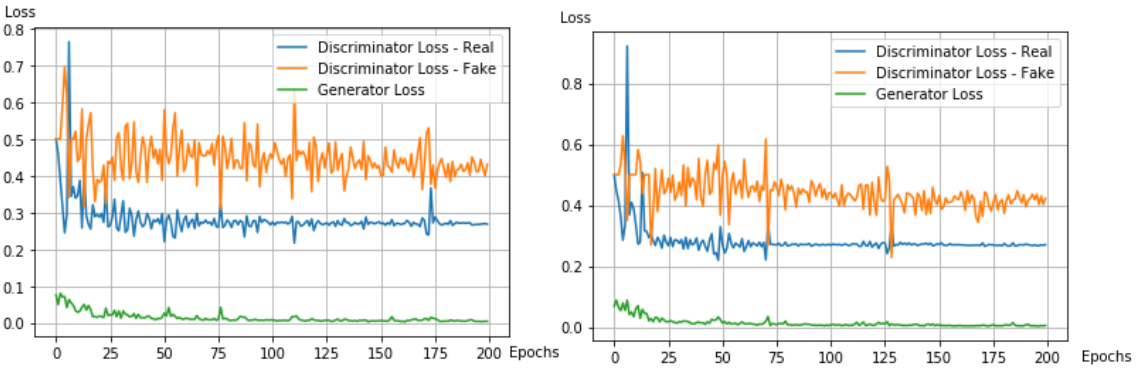

Figure 3.2: Convergence graphs of GAN model losses of training dataset. Left graph represents training on missing pixels, right - training on missing projetions.

Additionally, quantitative PSNR, SSIM and MSE metrics were calculated using validation dataset for both CNN and GAN, for further evaluation.

Missing counts		PSNR(dB)	SSMI	MSE
Reference (-10%)		31.3	0.673	4.521*10-6
Predicted (-10%)	CNN	45.67	0.993	0.016*10-6
	GAN	48.57	0.996	0.011*10-6
Reference (-30%)		31.05	0.627	1.080*10-5
Predicted (-30%)	CNN	45.11	0.992	0.002*10-5
	GAN	45.78	0.993	0.002*10-5
Reference (-50%)		30.4	0.586	1.344*10-5
Predicted (-50%)	CNN	42.75	0.988	0.003*10-5
	GAN	43.39	0.99	0.003*10-5
Reference (-70%)		30.12	0.567	1.242*10-5
Predicted (-70%)	CNN	40.64	0.982	0.059*10-5
	GAN	40.6	0.984	0.006*10-5
Reference (-90%)		29.09	0.554	0.771*10-5
Predicted (-90%)	CNN	36.11	0.952	0.02*10-5
	GAN	37.82	0.971	0.016*10-5

a)

Missing projectios		PSNR(dB)	SSMI	MSE
Reference (-10%)		39.43	0.967	0.057*10-6
Predicted (-10%)	CNN	45.35	0.993	0.018*10-6
	GAN	53.98	0.998	0.002*10-6
Reference (-30%)		28	0.9158	0.464*10-6
Predicted (-30%)	CNN	45.31	0.993	0.019*10-6
	GAN	52.63	0.997	0.003*10-6
Reference (-50%)		22.34	0.855	0.101*10-5
Predicted (-50%)	CNN	45.27	0.993	0.002*10-5
	GAN	53	0.997	0.0002*10-5
Reference (-70%)		14.59	0.575	0.218*10-5
Predicted (-70%)	CNN	41.27	0.981	0005*10-5
	GAN	45.58	0.994	0.002*10-5
Reference (-90%)		6.9	0.252	0.331*10-5
Predicted (-90%)	CNN	30.57	0.575	0.061*10-5
	GAN	38.52	0.98	0.0112*10-5

b)

Table 3.1: Table shows average PSNR, SSIM and MSE score for validation dataset. Table a) for data representation of missing pixels, and b) for missing projections. "Reference" stands for the corrupted data representation.

The green color on the table indicates which of the approaches outperformed the other on that particular data corruption level.

3.2 Quantitative results of test set

Test set contained 1 335 PET sinograms. This data set was put aside at the beginning of the project and left untouched until the final phase after the last models tuning and training were done to prevent any possible data leakage. The final results are presented in Table 3.2.

Missing counts		PSNR(dB)	SSMI	MSE
Reference (-10%)		30.39	0.644	0.533*10^-5
Predicted (-10%)	CNN	45.67	0.992	0.002*10^5
	GAN	47.44	0.994	0.001*10^-5
Reference (-30%)		30.27	0.601	1.281*10^-5
Predicted (-30%)	CNN	44.11	0.989	0.031*10^-5
	GAN	44.8	0.991	0.0266*10^-5
Reference (-50%)		28.84	0.526	1.601*10^-5
Predicted (-50%)	CNN	41.91	0.985	0.006*10^-5
	GAN	42.57	0.988	0.005*10^-5
Reference (-70%)		28.99	0.527	1.493*10^-5
Predicted (-70%)	CNN	40.1	0.979	0.009*10-5
	GAN	39.98	0.981	0.0094*10^-5
Reference (-90%)		27.05	0.482	0.995*10^-5
Predicted (-90%)	CNN	35.43	0.941	0.027*10-5
	GAN	36.91	0.963	0.023*10-5

a)

Missing projectios		PSNR(dB)	SSMI	MSE
Reference (-10%)		39.24	0.966	0.007*10^-5
Predicted (-10%)	CNN	44.56	0.941	0.002*10^-5
	GAN	51.74	0.996	0.005*10^-6
Reference (-30%)		27.64	0.914	0.061*10^-5
Predicted (-30%)	CNN	44.35	0.989	0.003*10^-5
	GAN	50.59	0.996	0.007*10^-6
Reference (-50%)		22.08	0.856	0.133*10^-5
Predicted (-50%)	CNN	44.51	0.992	0.003*10^-5
	GAN	50.94	0.996	0.006*10^-6
Reference (-70%)		14.89	0.534	0.267*10^-5
Predicted (-70%)	CNN	36.46	0.814	0.022*10^-5
	GAN	43.5	0.98402	0.004*10^-5
Reference (-90%)		6.59	0.243	0.432*10^-5
Predicted (-90%)	CNN	30.79	0.580569	0.082*10^-5
	GAN	37.44	0.971874	0.021*10^-5

b)

Table 3.2: Table shows average PSNR, SSIM and MSE score for test dataset. Table a) for data representation of missing pixels, and b) for missing projections. "Reference" stands for the corrupted data representation.

By comparing PSNR, GAN model (green color) outperformed CNN by scoring higher in all data corruption level of both data representations. Exceptionally, when predicting missing pixels with 70% of removed data information, CNN scored barely higher then GAN. SSIM metric confirmed that GAN performed better since it scored higher in each case without any exceptions.

3.3 Qualitative results of test set

Qualitative results are necessary to inspect how the model performed visually. Due to a high number of different cases which represent qualitative results, the figures are moved to the Appendix section and can be observed there. In general, qualitative results revealed that reconstructed PET images by GAN were sharper and more detailed. The edges of the tissue were more preserved, also less artificial noise and artifacts were induced by GAN, contrary to CNN.

Bibliography

[1] Feng Shi, David S. Lalush, Jiayin Kang, Weili Lin, Yaozong Gao, and Dinggang Shen. Prediction of standard-dose brain PET image by using MRI and low-dose brain [18 F]FDG PET images . *Medical Physics*, 42(9):5301–5309, 2015. doi: 10.1118/1.4928400. URL `http://dx.doi.org/10.1118/1.4928400`.

[2] Selena Milicevic Sephton and Simon M. Ametamey. *Positron emission tomography agents*. 2013. ISBN 9781909453401. doi: 10.4155/EBO.12.504.

[3] Frederic H Fahey. Data Acquisition in PET Imaging. pages 39–50, 2002.

[4] Biting Yu, Yan Wang, David S. Lalush, Dinggang Shen, Jiliu Zhou, Weili Lin, Xi Wu, Luping Zhou, Lei Wang, and Chen Zu. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *NeuroImage*, 174(October 2017): 550–562, 2018. ISSN 10538119. doi: 10.1016/j.neuroimage.2018.03.045.

[5] Sicong Yu and Hamed Hamid Muhammed. Comparison of pre- and post-reconstruction denoising approaches in positron emission tomography. In *Proceedings of 2016 1st International Conference on Biomedical Engineering: Empowering Biomedical Technology for Better Future, IBIOMED 2016*, pages 1–6, 2017. ISBN 9781509041428. doi: 10.1109/IBIOMED.2016.7869821.

[6] T. Leroy Laurence. Dead pixel compensation in PET, 2016.

[7] H.W.A.M. de Jong, R. Boellaard, C. Knoess, M. Lenox, C. Michel, M. Casey, and A.A. Lammertsma. Correction Methods for Missing Data in Sinograms of the HRRT PET Scanner. *IEEE Transactions on Nuclear Science*, 50(5):1452–1456, 2003. ISSN 0018-9499. doi: 10.1109/tns.2003.817366.

[8] F. Hashimoto, H. Ohba, K. Ote, and H. Tsukada. Denoising of Dynamic Sinogram by Image Guided Filtering for Positron Emission Tomography. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2(6):541–548, 2018. ISSN 2469-7311. doi: 10.1109/trpms.2018.2869936.

[9] Kostadin Dabov, Student Member, Alessandro Foi, Vladimir Katkovnik, Karen Egiazarian, and Senior Member. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. 16(8):2080–2095, 2007. ISSN 1057-7149. doi: 10.1109/TIP.2007.901238.

[10] Antoni Buades. A non-local algorithm for image denoising. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, II(0):60–65, 2005. ISSN 1063-6919. doi: 10.1109/CVPR.2005.38.

[11] Sari Peltonen, Uygar Tuna, Enrique Sánchez-Monge, and Ulla Ruotsalainen. PET sinogram denoising by block-matching and 3D filtering. *IEEE Nuclear Science Symposium Conference Record*, pages 3125–3129, 2012. ISSN 10957863. doi: 10.1109/NSSMIC.2011.6152568.

[12] M M Hasan. Adaptive Edge-Guided Block-Matching and 3D Filtering (BM3D) Image Denoising Algorithm. (May):1–120, 2014. URL papers://b6c7d293-c492-48a4-91d5-8fae456be1fa/Paper/p12096{%}5Cnfile:///C:/Users/Serguei/OneDrive/Documents/Papers/AdaptiveEdge-GuidedBlock-Matchingand3D-2014-04-30-1.pdf.

[13] Ulas Bagci and Daniel J. Mollura. Denoising PET images using singular value thresholding and Stein's unbiased risk estimate. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8151 LNCS(PART 3):115–122, 2013. ISSN 03029743. doi: 10.1007/978-3-642-40760-4_15.

[14] D. Visvikis, C. Cheze Le Rest, A. Le Pogam, H. Hanzouli, and M. Hatt. Denoising of PET images by combining wavelets and curvelets for improved preservation of resolution and quantitation. *Medical Image Analysis*, 17(8):877–891, 2013. ISSN 13618415. doi: 10.1016/j.media.2013.05.005. URL http://dx.doi.org/10.1016/j.media.2013.05.005.

[15] Leticia Ortega Maynez, Humberto de Jesus Ochoa Dominguez, Boris Mederos, Jose M. Mejia, and Osslan Osiris Vergara Villegas. Noise Reduction in Small-Animal PET Images Using a Multiresolution Transform. *IEEE Transactions on Medical Imaging*, 33 (10):2010–2019, 2014. ISSN 0278-0062. doi: 10.1109/tmi.2014.2329702.

[16] Yan Wang, Pei Zhang, Le An, Guangkai Ma, Jiayin Kang, Feng Shi, Xi Wu, Jiliu Zhou, David S. Lalush, Weili Lin, and Dinggang Shen. Predicting standard-dose PET image from low-dose PET and multimodal MR images using mapping-based sparse representation. *Physics in Medicine and Biology*, 61(2):791–812, 2016. ISSN 13616560. doi: 10.1088/0031-9155/61/2/791.

[17] Feng Shi, Le An, David S. Lalush, Ehsan Adeli, Pei Zhang, Weili Lin, Guangkai Ma, Yan Wang, and Dinggang Shen. Multi-Level Canonical Correlation Analysis for Standard-Dose PET Image Estimation. *IEEE Transactions on Image Processing*, 25(7):3303–3315, 2016. ISSN 1057-7149. doi: 10.1109/tip.2016.2567072.

[18] Habibi Heravi Hamed and Jahani Aghdam Elnaz. *Guide to convolutional neural networks : a practical application to traffic-sign detection and classification.* Cham, Switzerland : Springer, 2017. ISBN 9783319575506 3319575503.

[19] Le An, Dinggang Shen, Lei Xiang, Qian Wang, Dong Nie, Yu Qiao, and Weili Lin. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. *Neurocomputing*, 267:406–416, 2017. ISSN 09252312. doi: 10.1016/j.neucom.2017.06.048.

[20] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. pages 1–9, 2014. URL http://arxiv.org/abs/1406.2661.

[21] Phillip Isola, Jun Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:5967–5976, 2017. doi: 10.1109/CVPR.2017.632.

[22] Ji Zhao, Zhiqiang Chen, Li Zhang, and Xin Jin. Unsupervised Learnable Sinogram Inpainting Network (SIN) for Limited Angle CT reconstruction. (Dl):1–10, 2018. URL http://arxiv.org/abs/1811.03911.

[23] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, Lawrence Tarbox, and Fred Prior. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6):1045–1057, 2013. ISSN 08971889. doi: 10.1007/s10278-013-9622-7.

[24] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*, volume 13. MIT Press, 2017. ISBN 9780521835688. doi: 10.1038/nmeth.3707. URL http://goodfeli.github.io/dlbook/{%}0Ahttp://dx.doi.org/10.1038/nature14539.

[25] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Physics in Medicine and Biology*, 60(14):5471–5496, 2015. ISSN 13616560. doi: 10.1088/0031-9155/60/14/5471.

[26] TCIA. Soft-tissue-Sarcoma. URL https://wiki.cancerimagingarchive.net/display/Public/Soft-tissue-Sarcoma.

[27] Attenuation Correction and Positron Emission. Attenuation Correction, 1992. URL http://www.turkupetcentre.net/petanalysis/image{_}reconstruction.html.

[28] J Johansson. Radon transform in PET. 6(2):103, 2004.

[29] Brendan F. Hayden. Slice reconstrucion, 2005. URL http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL{_}COPIES/AV0405/HAYDEN/Slice{_}Reconstruction.html.

[30] Christopher M Bishop. *Pattern Recognition and Machine Learning*, volume 53. 2013. ISBN 978-0-387-31073-2. doi: 10.1117/1.2819119.

[31] Warren S. Sarle. Neural nets. URL url:ftp://ftp.sas.com/pub/neural/FAQ2.html.

[32] Swarnendu Ghosh, Nibaran Das, and Mita Nasipuri. Reshaping inputs for convolutional neural network: Some common and uncommon methods. *Pattern Recognition*, 93:79–94, 2019. ISSN 00313203. doi: 10.1016/j.patcog.2019.04.009. URL https://doi.org/10.1016/j.patcog.2019.04.009.

[33] Pytorch. Pytorch. URL https://pytorch.org/docs/stable/index.html.

[34] Keras. Keras. URL https://keras.io/.

[35] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351:234–241, 2015. ISSN 16113349. doi: 10.1007/978-3-319-24574-4_28.

[36] Hongming Shan, Yi Zhang, Qingsong Yang, Uwe Kruger, Mannudeep K. Kalra, Ling Sun, Wenxiang Cong, and Ge Wang. 3-D Convolutional Encoder-Decoder Network for Low-Dose CT via Transfer Learning From a 2-D Trained Network. *IEEE Transactions on Medical Imaging*, 37(6):1522–1534, 2018. ISSN 1558254X. doi: 10.1109/TMI.2018.2832217.

[37] Ugur Demir and Gozde Unal. Patch-Based Image Inpainting with Generative Adversarial Networks. 2018. URL http://arxiv.org/abs/1803.07422.

[38] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. pages 1–15, 2014. URL http://arxiv.org/abs/1412.6980.

[39] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss Functions for Image Restoration With Neural Networks. *IEEE Transactions on Computational Imaging*, 3(1):47–57, 2016. doi: 10.1109/tci.2016.2644865.

[40] Peter Aundal and John Aasted. The Radon Transform Theory and Implementation Peter Toft Department of Mathematical Modelling Section for Digital Signal Processing Technical University of Denmark. 1996.

[41] Oleg S. Pianykh, Ksenia Pospelova, and Nick H. Kamboj. Modeling Human Perception of Image Quality. *Journal of Digital Imaging*, 31(6):768–775, 2018. ISSN 1618727X. doi: 10.1007/s10278-018-0096-5.

[42] Wikipedia. Peak signal-to-noise ratio, 2011. URL http://en.wikipedia.org/wiki/Peak{_}signal-to-noise{_}ratio.

[43] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. ISSN 10577149. doi: 10.1109/TIP.2003.819861.

Appendix A

Appendix

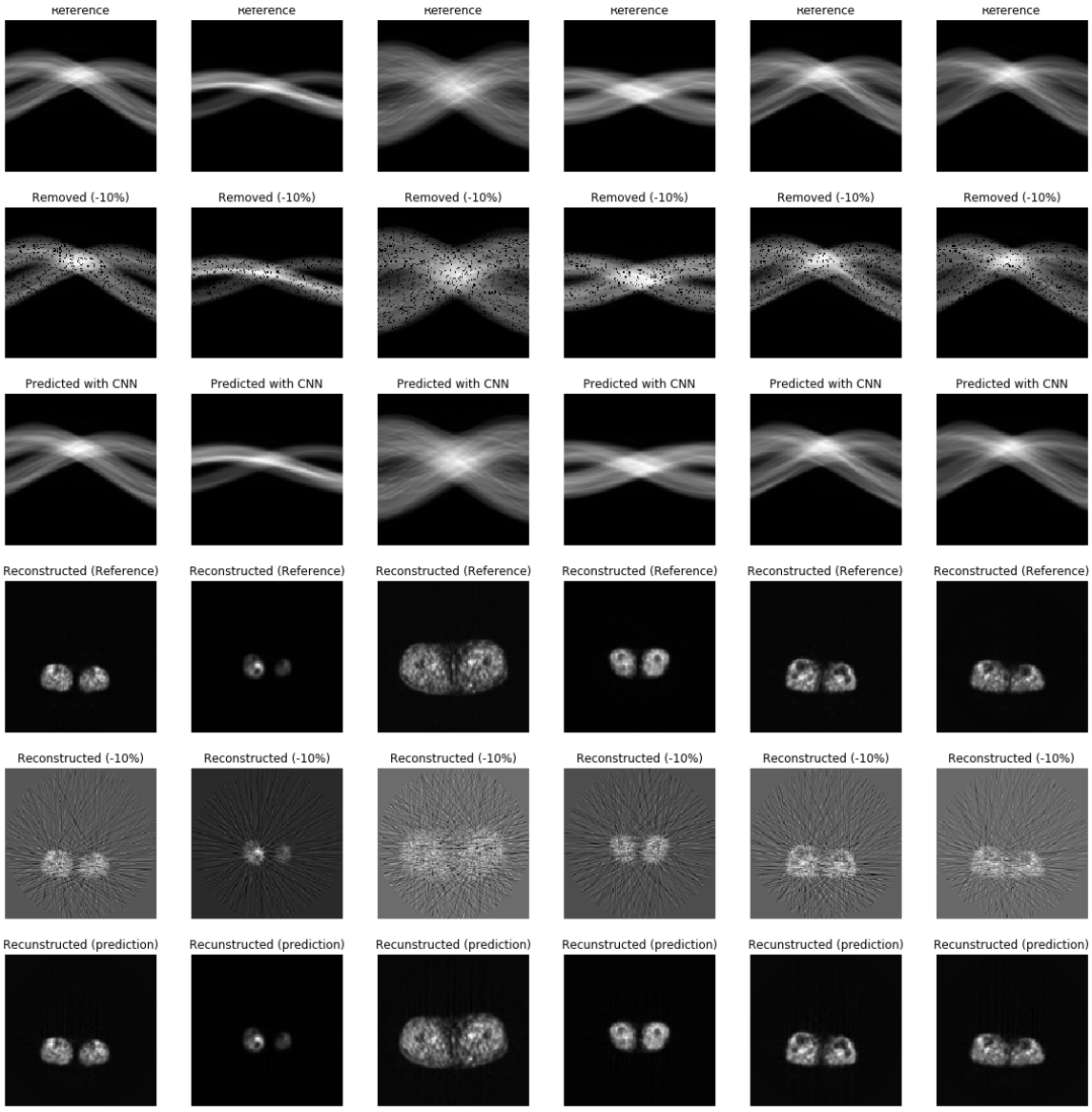

Figure A.1: Final results of CNN on test set when 10% of pixels are missing.

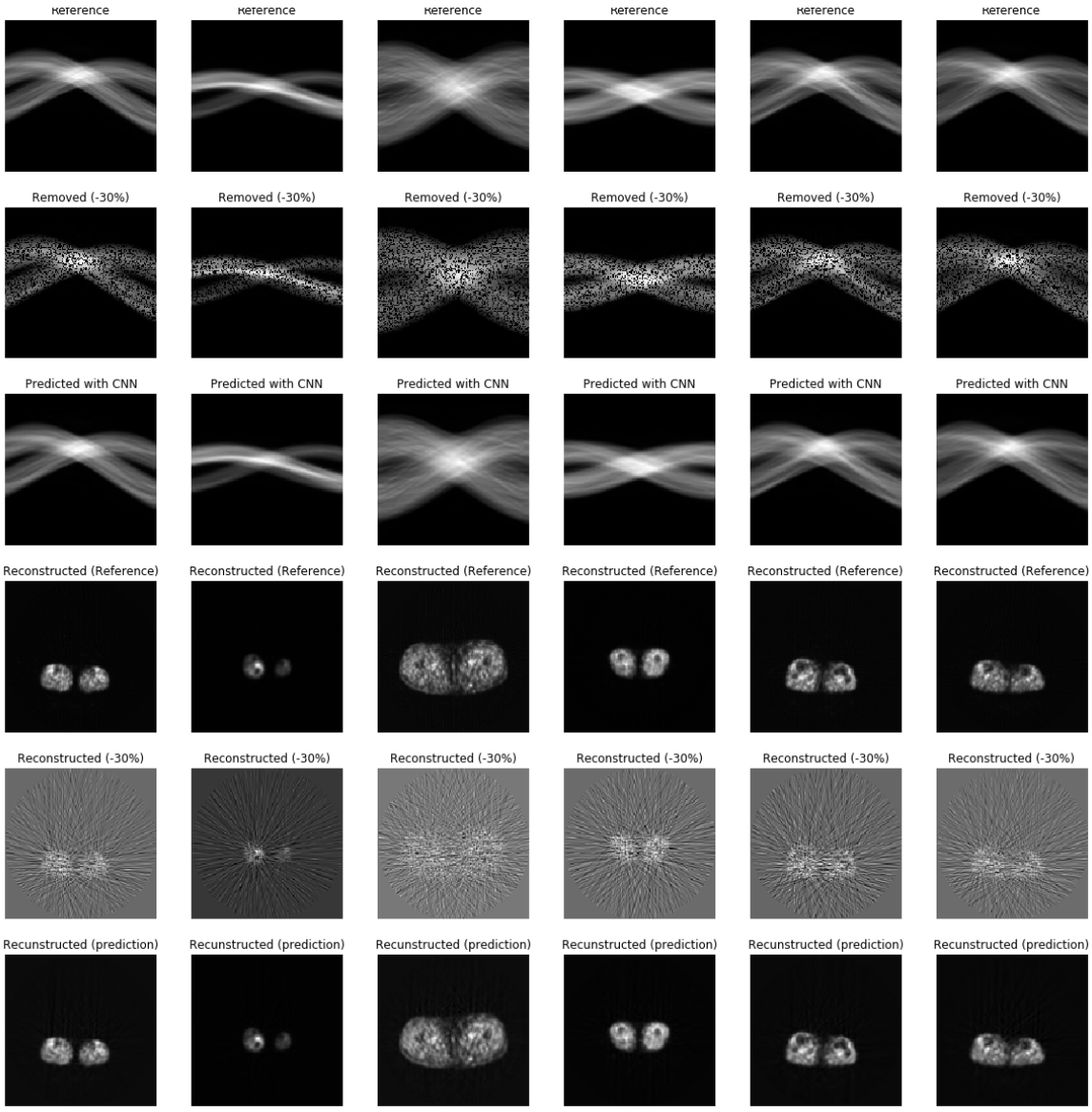

Figure A.2: Final results of CNN on test set when 30% of pixels are missing.

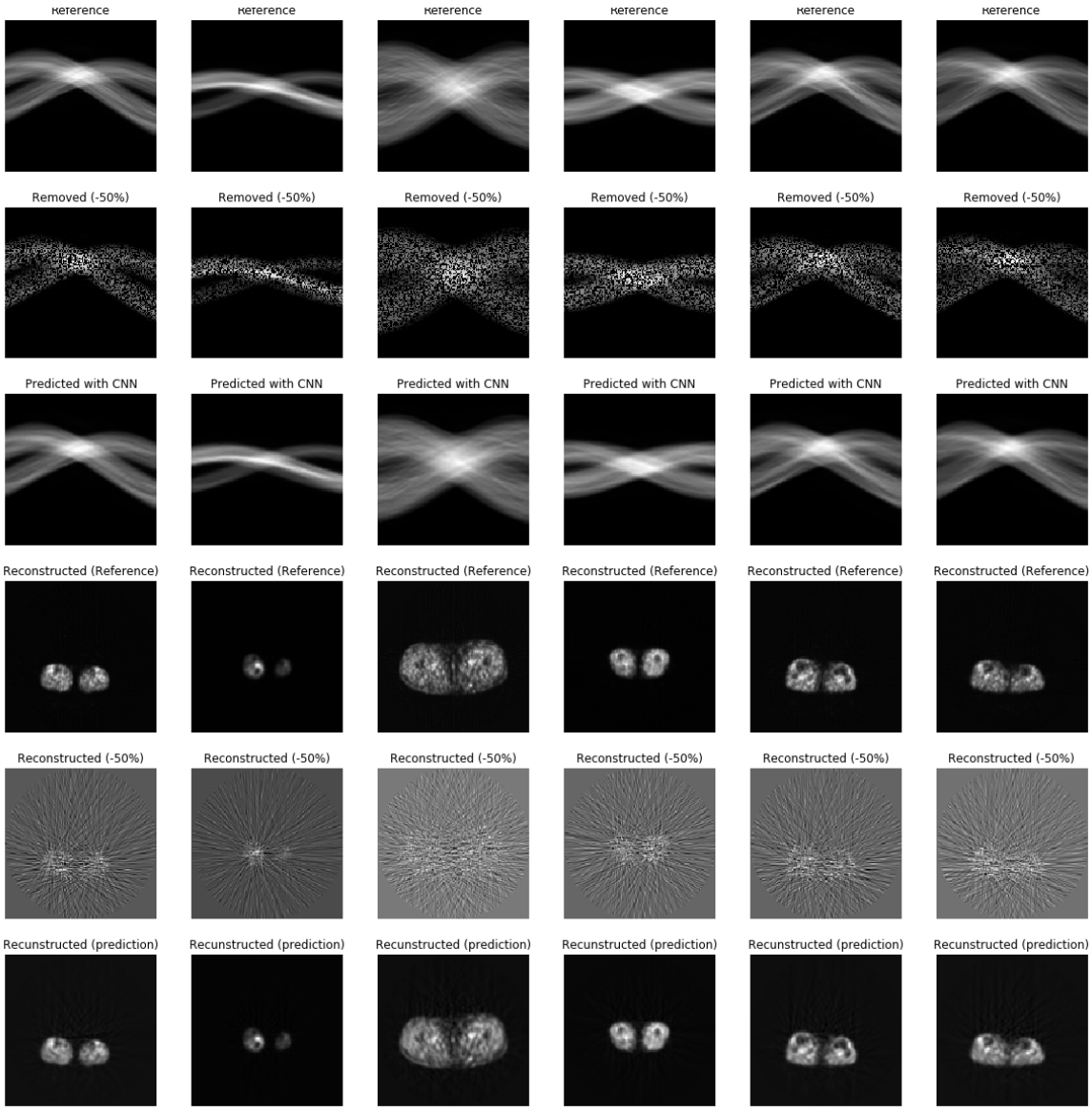

Figure A.3: Final results of CNN on test set when 50% of pixels are missing.

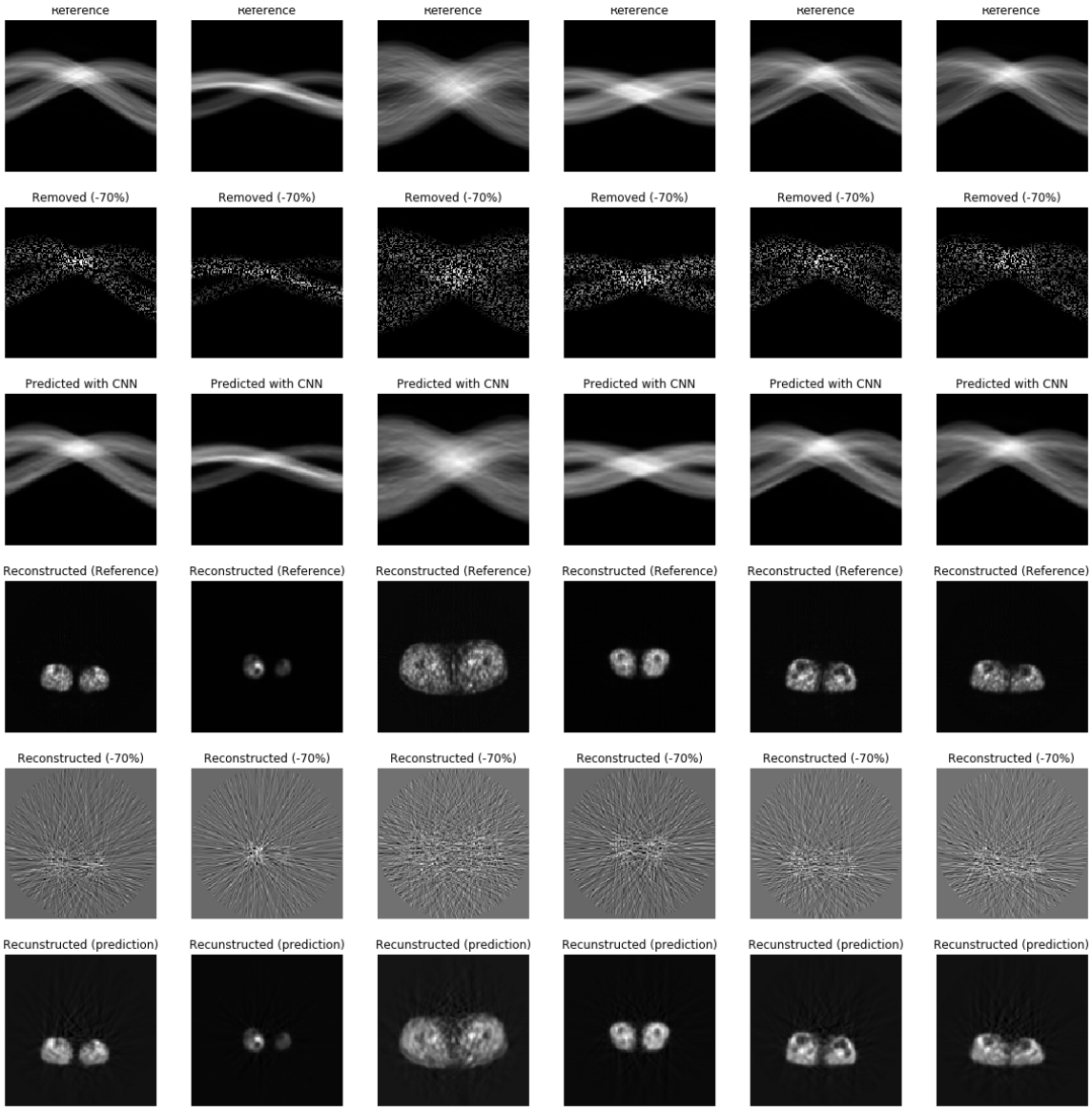

Figure A.4: Final results of CNN on test set when 70% of pixels are missing.

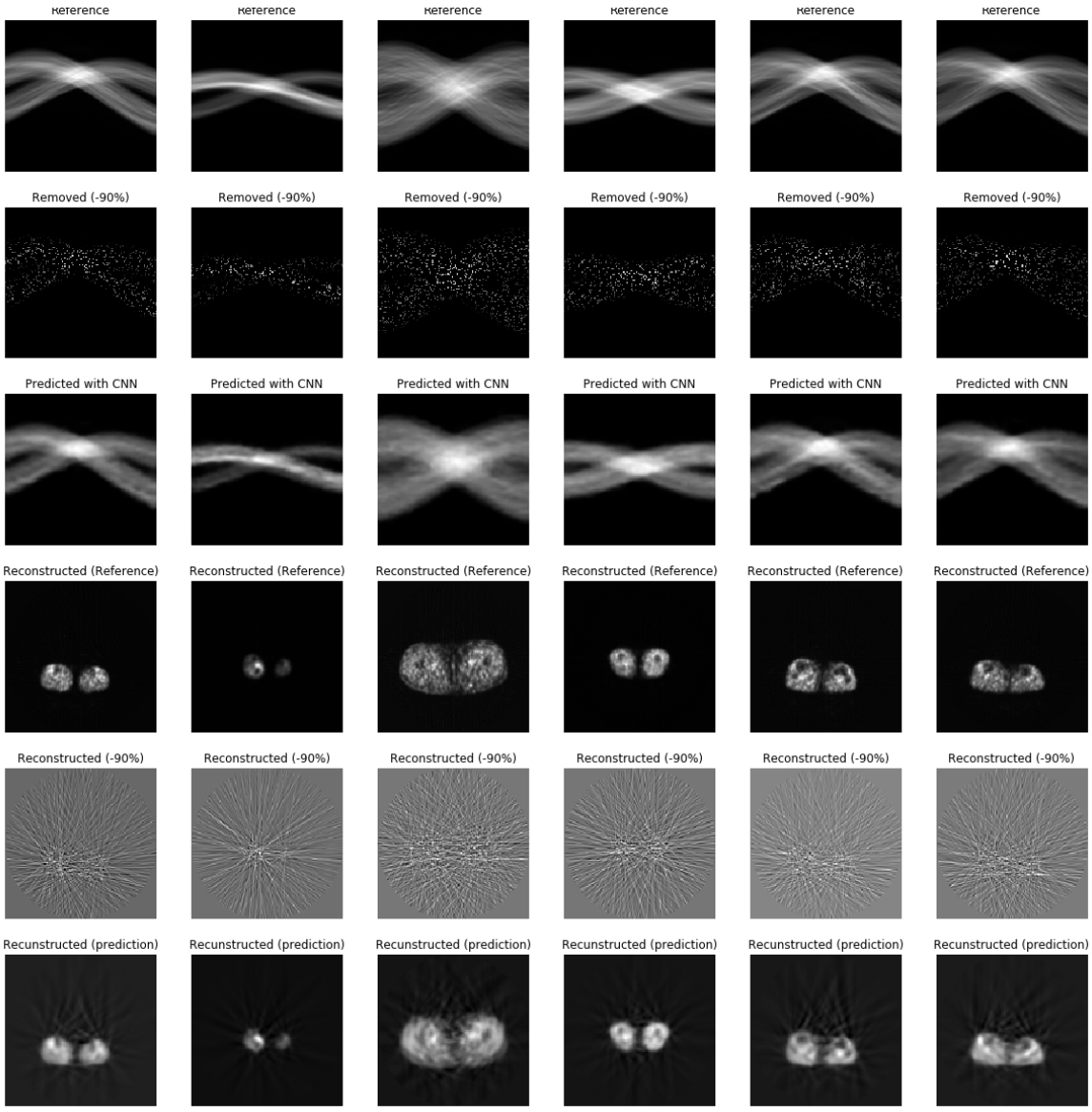

Figure A.5: Final results of CNN on test set when 90% of pixels are missing.

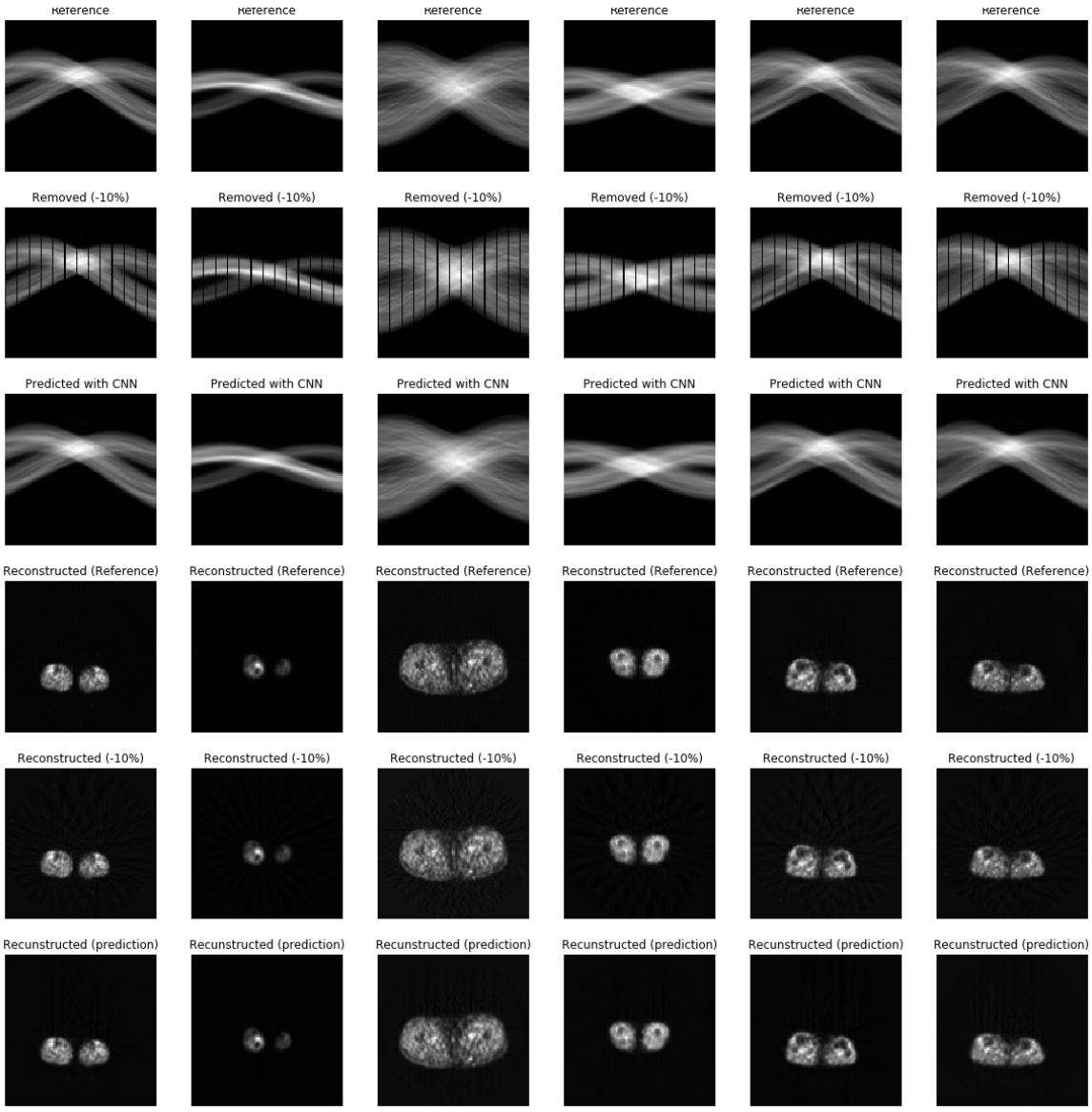

Figure A.6: Final results of CNN on test set when 10% of projections are missing.

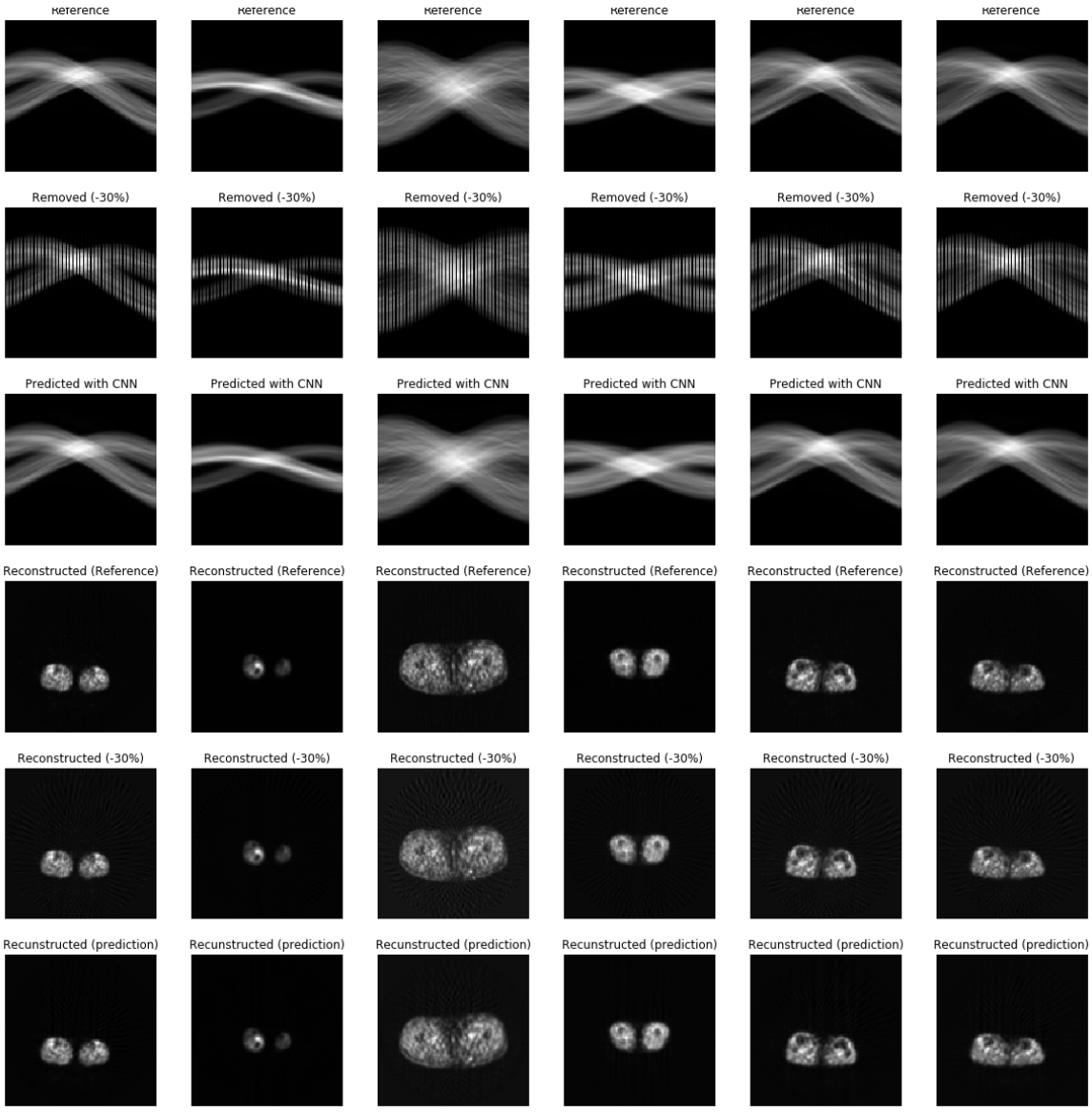

Figure A.7: Final results of CNN on test set when 30% of projections are missing.

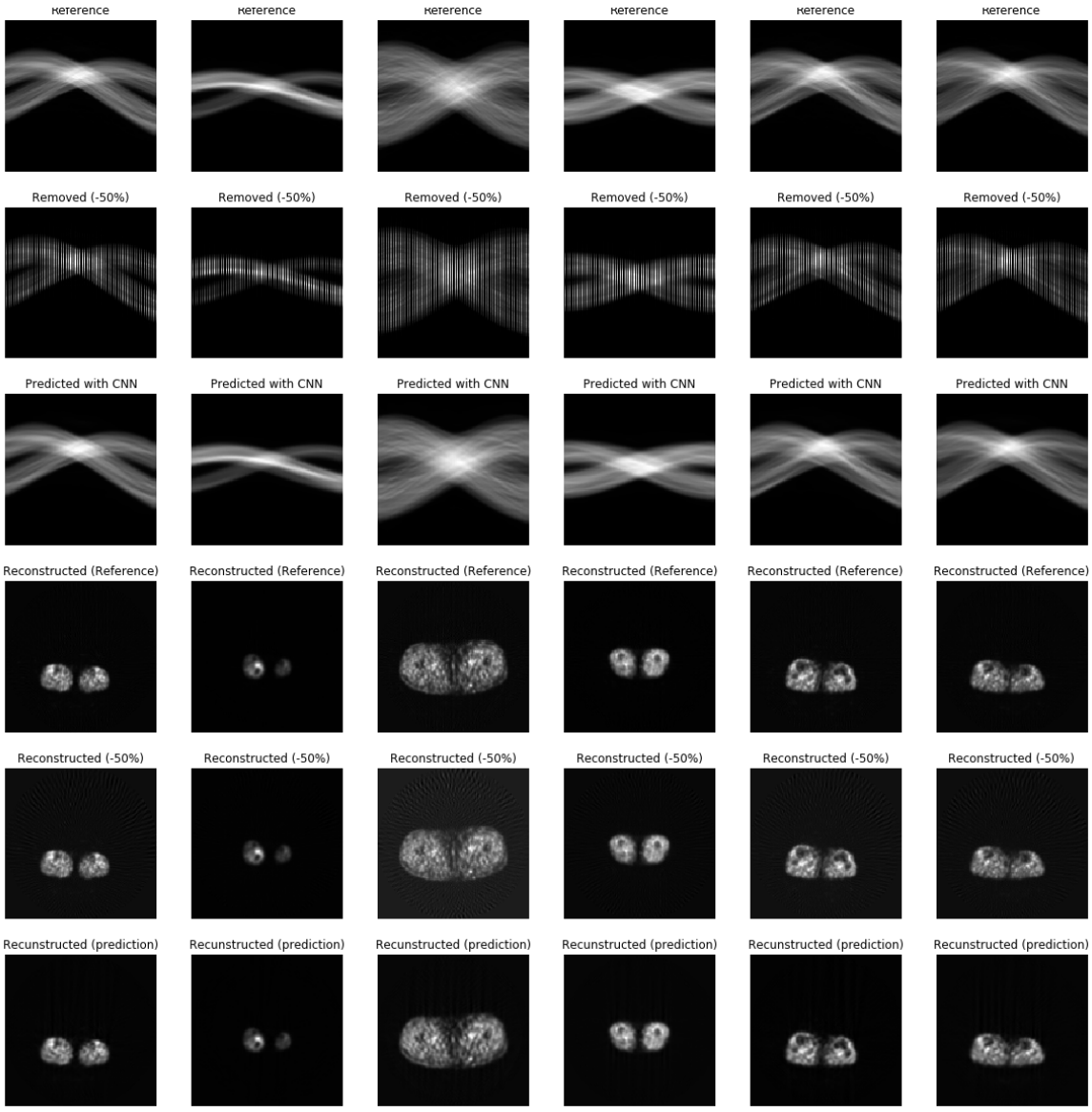

Figure A.8: Final results of CNN on test set when 50% of projections are missing.

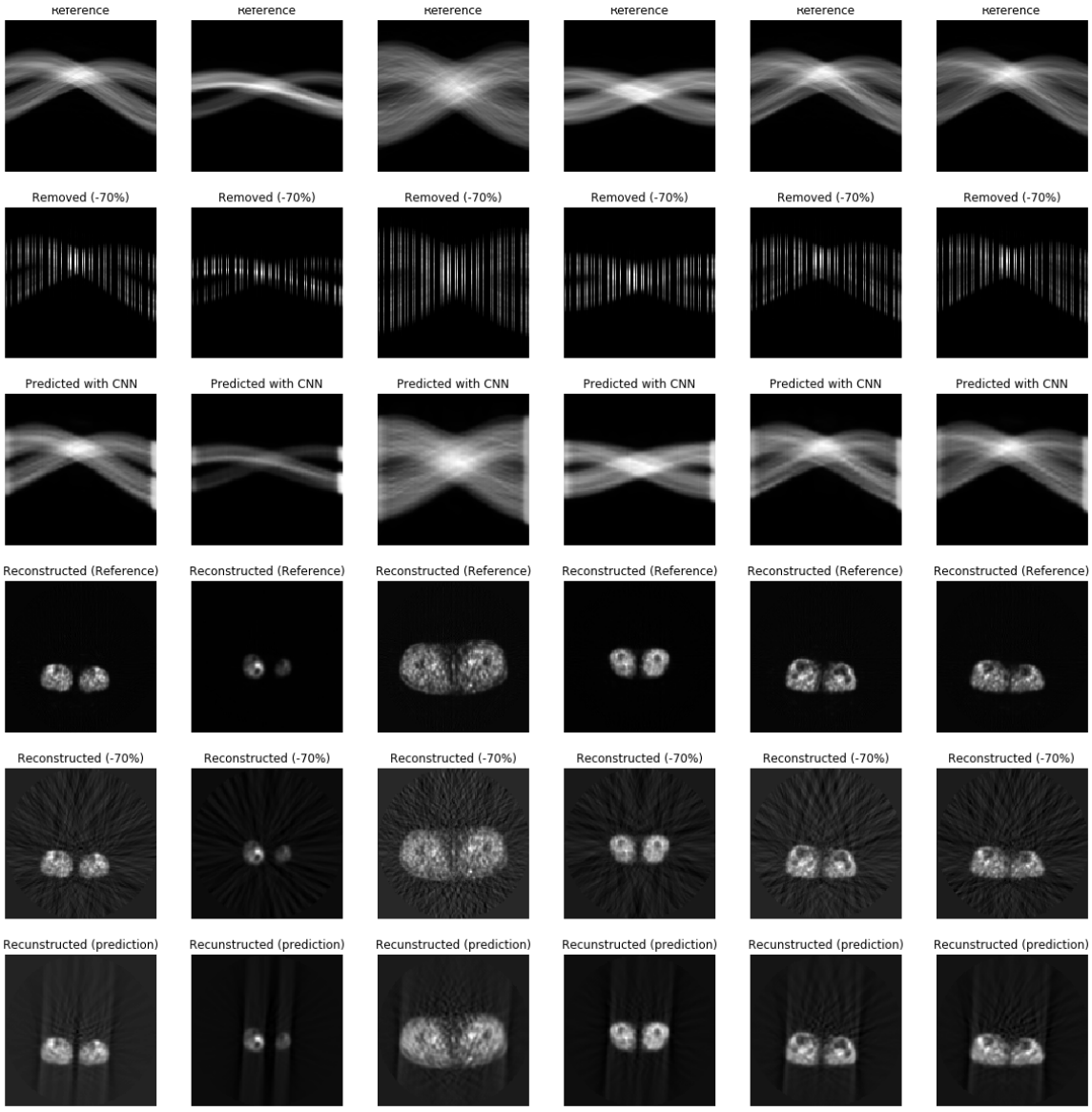

Figure A.9: Final results of CNN on test set when 70% of projections are missing.

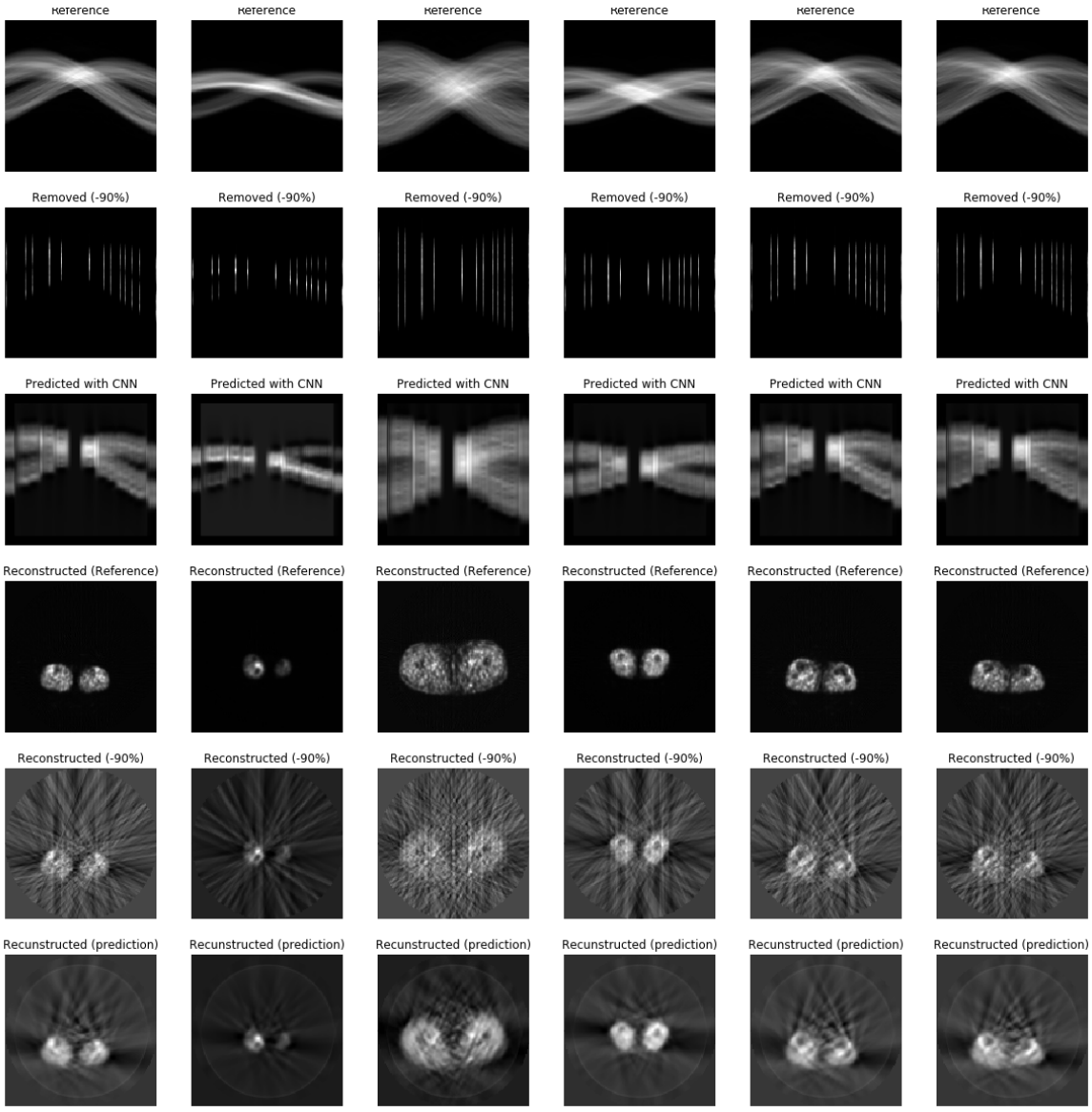

Figure A.10: Final results of CNN on test set when 90% of projections are missing.