

# Trajectory Upsampling for Sparse Conebeam Projections using Convolutional Neural Networks

Philipp Ernst<sup>1,2</sup>, Marko Rak<sup>1,2</sup>, Christian Hansen<sup>1,2</sup>, Georg Rose<sup>2,3</sup>, and Andreas Nürnberger<sup>1</sup>

<sup>1</sup>Faculty of Computer Science, University of Magdeburg, Germany

<sup>2</sup>Research Campus *STIMULATE*, University of Magdeburg, Germany

<sup>3</sup>Institute for Medical Engineering, University of Magdeburg, Germany

**Abstract** In this paper, we present an approach based on a combination of convolutional neural networks and analytical algorithms to interpolate between neighboring conebeam projections for upsampling along circular trajectories. More precisely, networks are trained to interpolate the angularly centered projection between the input projections of different angular distances. Experiments show that an analytical interpolation as additional input is more beneficial than adding more neighboring projections. Using our best model, we achieve an x8 upsampling by repeating the interpolation three times. Though not depending on a specific reconstruction algorithm, we show that FDK reconstructions substantially benefit from this upsampling for removing streak artifacts. Using this FDK reconstruction as initialization for ART is also superior to other initializations but comes with a higher computation time and therefore cannot be considered as an option in an interventional setting.

## 1 Introduction

Conebeam X-ray CT (CBCT) is a helpful tool for surgeons to guide them during interventions. The downside is that the patients as well as the surgeons are exposed to harmful X-radiation. Reducing it is possible by acquiring fewer projections or applying less radiation while keeping the number of projections high. In both cases however, the image quality of the reconstructed volumes using the commonly used FDK algorithm is severely impaired by streak artifacts or noise.

Many algorithms have been proposed to overcome these artifacts [1] but are too costly in terms of computation time to be applicable in an interventional setting, especially iterative methods that need to compute both forward and backprojections in each iteration for the entire 3-D volume.

Convolutional neural networks (CNNs) and deep learning have found their way into medical imaging [2] and CT image reconstruction [3]. Owing to their trainable nature, they directly incorporate domain knowledge to approximate the reconstruction more closely. Despite the high computational power of modern PC systems, training CNNs on whole 3-D data sets is usually not feasible due to the large memory requirements and is often stripped down to 2-D problems or patch-based 3-D approaches. The inherent two-dimensionality of conebeam projections suggests using them in combination with CNNs. The method proposed here will be used to interpolate between these projections which enables an upsampling along a circular trajectory around the scanned subject. Since the interpolation is carried out in projection space, the method does not rely on any reconstruction algorithm and preserves data consistency.

## 2 Method

### 2.1 Analytical Projection Interpolation

As described in [4], conebeam projections can be approximately interpolated by using (Eq. 24 in [4])

$$g(\lambda + \varepsilon\Delta\lambda, \underline{\alpha}) \simeq (1 - \varepsilon)g(\lambda, \underline{b}(\lambda + \varepsilon\Delta\lambda, \underline{\alpha}) - \underline{a}(\lambda)) + \varepsilon g(\lambda + \varepsilon\Delta\lambda, \underline{b}(\lambda + \varepsilon\Delta\lambda, \underline{\alpha}) - \underline{a}(\lambda + \Delta\lambda)) \quad (1)$$

for projections  $g(\lambda, \underline{\alpha})$  from source positions  $\underline{a}(\lambda)$  in directions  $\underline{\alpha}$  and points of interest  $\underline{b}(\lambda, \underline{\alpha})$  that are closest to the rotation axis on the line through  $\underline{a}(\lambda)$  with direction  $\underline{\alpha}$ . Unlike [4], the directions  $\underline{\alpha}$  here are chosen to coincide with the projection lines of the projection to be interpolated. This only requires interpolating on the given projections.

### 2.2 CNN Approach

Assuming an equiangular sampling of conebeam projections along a circular trajectory, the presented approach upsamples along the trajectory by subsequently interpolating projections angularly centered between neighboring projections. Simple algorithms like linear interpolation are not applicable because of the sinusoidal structure and perspective distortions caused by the conebeam. A U-Net [5] is used to approximate this highly complex interpolation because of its large receptive field that is able to capture and trace larger translations in the projections compared to flat CNN architectures. (1) Networks are trained to predict the projection angularly centered between two projections from only its direct neighbors for 2°, 4° and 8° of angular distance (referred to as nn2). (2) The number of neighboring input projections is increased from 2 to 4 and 8 neighbors to provide more angular information (referred to as nn4, nn8). (3) Instead of increasing the number of neighboring projections, the analytical interpolation described in Sec. 2.1 with  $\varepsilon = 0.5$  is used as an additional input which is supposed to guide the network closer to the true interpolation (referred to as nn2+ana).

### 2.3 Datasets and Training

The data of 22 subjects from the CT Lymph Nodes collection [6] of The Cancer Imaging Archive [7] is used, consisting of reconstructed volumes of the abdomen with different in-plane spacings that serve as ground truth. Conebeam projec-

Up	Method	NMSE ( $\times 10^{-5}$ )	PSNR [dB]	SSIM [%]
x2	ana	10.88	98.45	<b>99.26</b>
x2	nn2	<b>8.63</b>	<b>98.91</b>	99.01
x2	nn4	11.49	97.85	98.80
x2	nn8	17.32	95.74	97.87
x2	nn2+ana	10.91	97.97	98.74
x4	ana	32.60	93.65	97.84
x4	nn2	<b>17.15</b>	<b>95.73</b>	<b>98.24</b>
x4	nn4	27.63	93.69	97.53
x4	nn8	40.20	92.27	96.34
x4	nn2+ana	18.24	95.61	98.01
x8	ana	93.52	89.26	94.44
x8	nn2	58.64	90.87	96.03
x8	nn4	79.94	89.49	94.39
x8	nn8	114.83	87.93	92.61
x8	nn2+ana	<b>32.45</b>	<b>93.34</b>	<b>96.98</b>

**Table 1:** Projection errors for different upsampling methods.

tions were generated using the CTL toolkit [8] equiangularly along a circular trajectory with a source to detector distance (SDD) of 1000 mm and a source to isocenter distance (SID) of 750 mm. The flat panel detector consists of  $256 \times 256$  elements with a pixel size of  $4 \text{ mm}^2$  (cone angle of  $54.2^\circ$ ). The values were chosen such that most projections were not truncated and to enable a faster training.

The U-Net [5] has a depth of 5 and is slightly modified. The encoder doubles the number of layers after each average pooling, whereas the decoder halves the number of layers after each nearest neighbor upsampling. The optimizer is SGD with a weight decay of  $1 \times 10^{-4}$  and a learning rate of  $6 \times 10^{-3}$  that gradually drops to  $1 \times 10^{-6}$  by a factor of 0.8 after every 10 epochs of no improvement in validation loss. Every network was trained for 300 epochs using mean squared error (MSE) and another 300 epochs using equally weighted  $l_1$  and MS-SSIM loss similar to [9] to focus more on general structures and edges. 16, 4 and 2 datasets were used for training, validation and testing, respectively. For faster convergence, the projections were normalized between 0 and approximately 1 by dividing by the 99th percentile of all projections of all datasets.

### 3 Results

#### 3.1 Projections

The different interpolation methods are evaluated on the projections first. Except for the analytical upsampling described in Sec. 2.1, all methods interpolate the projection angularly centered between the input projections, which is repeated for x4 and x8 upsampling using the corresponding trained networks. For the analytical upsampling, the parameter  $\varepsilon$  is

Method	NMSE [%]	PSNR [dB]	SSIM [%]
full	4.95	82.95	99.14
sparse	16.09	75.05	97.72
ana	7.43	79.22	98.81
nn2	6.51	80.26	98.97
nn4	6.74	79.96	98.92
nn8	7.28	79.40	98.83
nn2+ana	<b>6.00</b>	<b>81.00</b>	<b>99.03</b>

**Table 2:** Reconstruction errors of FDK reconstructions for different upsampling methods from 45 available projections.

chosen to directly resemble the positions of the projections to be interpolated. Tab. 1 shows the results for the error metrics normalized mean squared error (NMSE), peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) averaged over all projections. The calculation of the metrics obviously excludes the non-interpolated projections. Interestingly, the results are quite different for the different upsampling stages.

For the single interpolation (x2, angular difference of  $2^\circ$ ), nn2 gives the best results for NMSE and PSNR. The analytical interpolation however results in the highest SSIM.

Interpolating twice (x4, angular difference of  $4^\circ$ ) is done best by nn2, this time for all metrics.

Finally, the optimal method for carrying out the interpolation three times (x8, angular difference of  $8^\circ$ ) is using nn2+ana.

A patch of an exemplary x8 interpolation created with the different methods is shown in Fig. 1. Compared to the ground truth patch, the other patches are more blurry. The patch created with the analytical interpolation looks like the superimposition of two projections. The nn4 and nn8 patches seem to have more high frequencies than nn2 and consequently look less blurry. nn2+ana is visually closest to the ground truth and the least blurred.

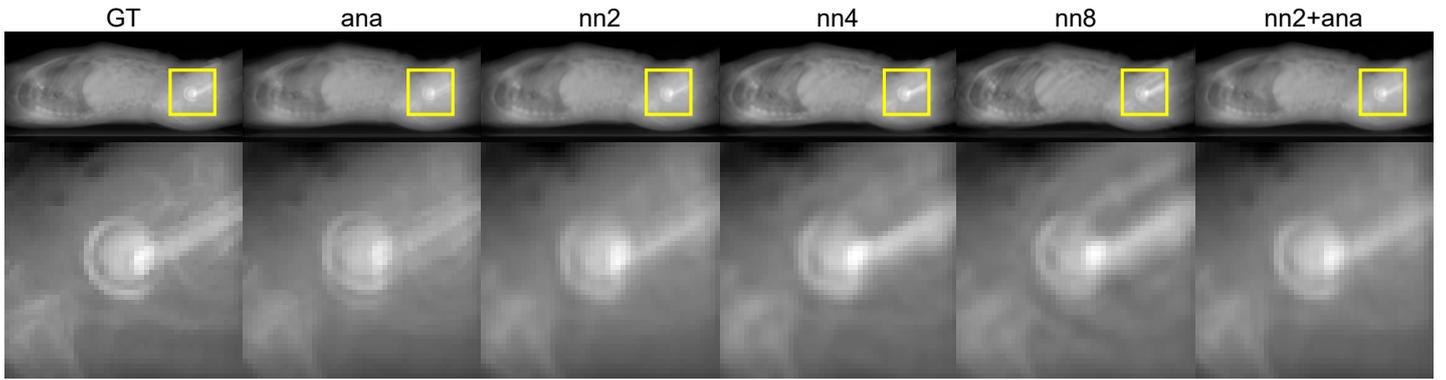
#### 3.2 Reconstructions

Evaluating in projection space only does not fully show the benefits of the proposed method. It is also necessary to compare the reconstructions. We decided for the commonly used FDK [10] algorithm as well as ART [11] without interpolated projections initialized with the FDK reconstruction using all interpolated projections.

All reconstructions are created with the CTL toolkit [8]. The ART reconstructions run for 5 iterations with enabled positivity constraint.

Since the number of projections is still relatively small and the resolution of the detector is quite low, the reconstructions will also be compared to the FDK reconstruction using all 360 projections to find lower bounds for the error metrics.

As described previously, though not depending on any recon-



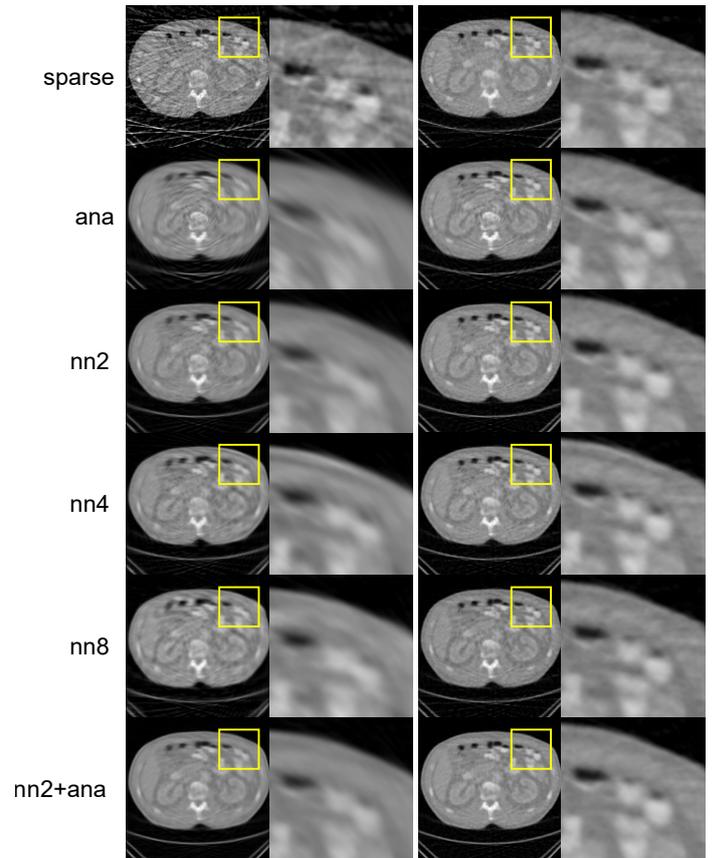
**Figure 1:** Top: Interpolated projections (central part) of x8 upsampling of different interpolation methods compared to ground truth projection (GT). Bottom: Zoomed patches around a hip implant.

struction algorithm, the interpolated projections are supposed to increase the quality of the reconstructions by providing a more appropriate sampling of projections.

This hypothesis is evaluated using the FDK reconstruction algorithm, first. For brevity, only the reconstructions of the highest upsampling (x8, 45 available projections) are investigated. Tab. 2 shows the error metrics for the different methods averaged over all axial slices. For reference, the first two rows serve as lower/upper bounds: values for the full FDK describe the errors between the ground truth volume and the volume reconstructed from 360 projections, whereas values for the sparse FDK describe the errors between ground truth and reconstruction from 45 projections. All interpolation methods optimize the sparse FDK reconstruction and are quantitatively closer to the full FDK. nn2+ana works best, followed by nn2, nn4, nn8 and using only the analytical interpolation. This closely resembles the errors on the projections described in the previous section.

The left column of Fig. 2 shows exemplary FDK reconstructions using the different methods. Compared to the direct FDK reconstruction from 45 projections (sparse), every method reduces the streak artifacts. The analytical upsampling (ana), however, basically results in a radially blurred reconstruction. None of the CNN-based reconstructions suffers from streak artifacts or radial blur, but they appear slightly more blurred than the sparse FDK reconstruction. As expected from the quantitative analysis, nn2+ana also creates the best visual result.

ART provides another simple reconstruction algorithm. Due to its iterative nature, it is inherently slower than FDK but enables simply adding additional constraints resulting in reconstructions of higher quality. For a better convergence, ART is initialized with another reconstruction. In our experiments, we use the FDK reconstructions of the different interpolation methods and run ART with only the 45 available projections, which results in the best compromise between reconstruction time and quality. Tab. 3 shows the error metrics. Zero-initialized ART and sparse-FDK-initialized ART are shown for reference. In all cases, ART outperforms FDK. Again, nn2+ana works best, followed by the other methods



**Figure 2:** Reconstructions for different upsampling methods. Left column: FDK. Right column: ART initialized with FDK.

Init.	NMSE [%]	PSNR [dB]	SSIM [%]
zero	2.72	82.35	99.73
sparse	2.28	83.10	99.78
ana	2.19	83.25	99.80
nn2	1.88	83.98	99.84
nn4	1.96	83.76	99.83
nn8	2.12	83.42	99.81
nn2+ana	<b>1.65</b>	<b>84.59</b>	<b>99.86</b>

**Table 3:** Reconstruction errors of ART reconstructions for different upsampling methods from 45 available projections.

in the same order as in the FDK reconstructions.

The right column of Fig. 2 shows exemplary ART reconstructions using the different methods. They are not only quantitatively closer to the ground truth but also qualitatively outperform their FDK counterparts. There are only slight visual differences of the ART initialized with the different FDK reconstructions. For the sparse case, edges are preserved well but tissues of the same absorption coefficient appear noisy. nn2+ana has the best visual quality with the least noise and the best edge preservation compared to the other methods.

## 4 Discussion

Increasing the number of neighboring projections does not increase the quality of the interpolated projections. Since the additional projections are only provided to the CNN as input channels and the convolutions are carried out per channel, it is possible that (without any special weight initialization) the information from more distant neighbors is not local enough to be considered as helpful knowledge during backpropagation. Moreover, increasing the number of input projections even impairs the prediction quality. Further tests need to investigate why different interpolation methods work best for certain upsampling stages.

The simulated projections do not contain noise, are almost not truncated, have a low resolution and a rather large pixel spacing. Further experiments need to focus on more realistic detector and gantry parameters and the method needs to be tested on real data, especially including interventional instruments and other artifact creating influences.

The used error metrics only give a rough impression of the quality. Due to the blurring of edges caused by the interpolation, future work needs to focus on how exactly mappings of edges are changed as well as how the reconstructions compare to other state-of-the-art methods.

Using the neighboring projections as input channels of the U-Net is a rather straightforward way. As with other deep learning methods, it is conceivable that another network architecture can extract more information from the input data and thus improve the quality even further, which will be part of future experiments. The code is available on Github<sup>1</sup>.

## 5 Conclusion

It was shown that conebeam projection interpolation using CNNs applied to trajectory upsampling significantly reduces streak artifacts from FDK reconstructions and provides a strong prior for iterative reconstruction algorithms when used for the initialization in ART. Providing further knowledge about the interpolation to the network in terms of the analytical interpolation approach similar to [4], the quality can be improved even further. This allows for a dose reduction by

a factor of at least eight while still providing a good quality of the reconstructions. Compared to an FDK reconstruction from 45 projections, our best interpolation method increases the PSNR by almost 6 dB. Though not applicable in interventions due to time requirements, initializing an ART with the FDK reconstruction further increases the PSNR to 84.59 dB.

## Acknowledgements

This work was conducted within the International Graduate School MEMoRIAL at OVGU Magdeburg, supported by the ESF (project no. ZS/2016/08/80646).

## References

- [1] M. Bertram, J. Wiegert, D. Schafer, et al. "Directional View Interpolation for Compensation of Sparse Angular Sampling in Cone-Beam CT". *IEEE Transactions on Medical Imaging* 28.7 (2009), pp. 1011–1022. DOI: [10.1109/TMI.2008.2011550](https://doi.org/10.1109/TMI.2008.2011550).
- [2] P. Ernst, G. Hille, C. Hansen, et al. "A CNN-Based Framework for Statistical Assessment of Spinal Shape and Curvature in Whole-Body MRI Images of Large Populations". *MICCAI 2019*. Springer International Publishing, 2019, pp. 3–11.
- [3] J. Adler and O. Öktem. "Learned Primal-Dual Reconstruction". *IEEE Transactions on Medical Imaging* 37.6 (2018), pp. 1322–1332. DOI: [10.1109/TMI.2018.2799231](https://doi.org/10.1109/TMI.2018.2799231).
- [4] F. Noo, S. Hoppe, F. Dennerlein, et al. "A new scheme for view-dependent data differentiation in fan-beam and cone-beam computed tomography". *Physics in Medicine and Biology* 52.17 (Aug. 2007), pp. 5393–5414. DOI: [10.1088/0031-9155/52/17/020](https://doi.org/10.1088/0031-9155/52/17/020).
- [5] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". *MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, et al. Cham: Springer International Publishing, 2015, pp. 234–241. DOI: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [6] H. R. Roth, L. Lu, A. Seff, et al. *A new 2.5 D representation for lymph node detection in CT. The Cancer Imaging Archive*. 2015. DOI: [10.7937/K9/TCIA.2015.AQI1DCNM](https://doi.org/10.7937/K9/TCIA.2015.AQI1DCNM).
- [7] K. Clark, B. Vendt, K. Smith, et al. "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository". *Journal of Digital Imaging* 26.6 (Dec. 2013), pp. 1045–1057. DOI: [10.1007/s10278-013-9622-7](https://doi.org/10.1007/s10278-013-9622-7).
- [8] T. Pfeiffer, R. Frysche, R. N. K. Bismark, et al. "CTL: modular open-source C++-library for CT-simulations". *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. Ed. by S. Matej and S. D. Metzler. Vol. 11072. International Society for Optics and Photonics. SPIE, 2019, pp. 269–273. DOI: [10.1117/12.2534517](https://doi.org/10.1117/12.2534517).
- [9] H. Zhao, O. Gallo, I. Frosio, et al. "Loss Functions for Image Restoration With Neural Networks". *IEEE Transactions on Computational Imaging* 3.1 (2017), pp. 47–57. DOI: [10.1109/TCI.2016.2644865](https://doi.org/10.1109/TCI.2016.2644865).
- [10] L. A. Feldkamp, L. C. Davis, and J. W. Kress. "Practical cone-beam algorithm". *J. Opt. Soc. Am. A* 1.6 (June 1984), pp. 612–619. DOI: [10.1364/JOSA.1.000612](https://doi.org/10.1364/JOSA.1.000612).
- [11] R. Gordon, R. Bender, and G. T. Herman. "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography". *Journal of Theoretical Biology* 29.3 (1970), pp. 471–481. DOI: [https://doi.org/10.1016/0022-5193\(70\)90109-8](https://doi.org/10.1016/0022-5193(70)90109-8).

<sup>1</sup>[https://github.com/phernst/conebeam\\_interpolation](https://github.com/phernst/conebeam_interpolation)