Differentiable Forward Projector for X-ray Computed Tomography

Hyojin Kim [*1] **Kyle Champley** [*2]

Abstract

Data-driven deep learning has been successfully applied to various computed tomographic reconstruction problems. The deep inference models may outperform existing analytical and iterative algorithms, especially in ill-posed CT reconstruction. However, those methods often predict images that do not agree with the measured projection data. This paper presents an accurate differentiable forward and back projection software library to ensure the consistency between the predicted images and the original measurements. The software library efficiently supports various projection geometry types while minimizing the GPU memory footprint requirement, which facilitates seamless integration with existing deep learning training and inference pipelines. The proposed software is available as open source: https://github.com/LLNL/LEAP.

1. Introduction

X-ray Computed Tomography (XCT) is a 3D noninvasive/nondestructive imaging modality that has applications in healthcare, security, and industry (H. E. Martz et al., 2016). X-ray projections are recorded by an x-ray detector opposite of an x-ray source over a range of angles. The XCT inverse problem attempts to reconstruct the 3D volume of x-ray linear attenuation coefficients from this collection of x-ray projections. The XCT reconstruction problem has been extensively studied for several decades. Recently, deep learning (DL)-based inference models have successfully tackled more challenging reconstruction problems, such as scenarios where x-ray projections are only collected over a small range of angles (limited-angle CT) or where few projections are collected (sparse-view or few-view CT) which are characterized as ill-posed inverse problems (Han & Ye, 2018; Zhou & Horstmeyer, 2020; Zhang et al., 2020; 2021; Gao et al., 2022; Hu et al., 2022).

While recent advancements in deep neural network trained models have been shown to outperform existing numerical optimization methods in some applications, they often suffer from a lack of data consistency to check whether the predicted images agree with the measured X-ray projection data (sinogram). Moreover, those inference models rely mainly on the training data distribution with the loss functions typically based on the reconstruction error between the predicted images and the ground-truth images. In these XCT inverse problems, it is crucial to incorporate the data consistency step into the inference models to suppress potential artifacts and hallucinated regions. To enable this, a proper forward model needs to be incorporated into the training pipeline or the inference step.

One way to facilitate this is the use of an additional data consistency step with a loss function in the training or inference pipeline. The loss functions typically aim to minimize the difference between the forward-projected data of the predicted volume and the original projection data (Zhou et al., 2021; Liu et al., 2022; Lahiri et al., 2023). In this setup, a differentiable forward projection model needs to be integrated into the neural network framework. Another way is to use the predicted volumes from the inference models as a prior (initial guess) in additional reconstruction algorithms (Kim et al., 2019; Anirudh et al., 2019; Zhou & Horstmeyer, 2020). Similarly, the inferred volumes are forward projected to complete the projection data in limited-angle CT applications (Anirudh et al., 2018). Although the additional reconstruction algorithms and the sinogram completion step can be performed separately, an integrated end-to-end pipeline including the neural network inference models and following reconstruction processes make the application more cohesive and practical.

Despite the increased demand for differentiable projectors in a wide range of XCT applications, existing solutions for differentiable projections are somewhat limited and impractical. The Radon Transform is widely used, but works only with the parallel beam geometry. One could pre-compute and store the projection matrix for any desired CT geometry

[*]Equal contribution [1]Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, California, USA [2]Ziteo Medical, California, USA, He conducted the software development and experiments while affiliated with Lawrence Livermore National Laboratory. Correspondence to: Hyojin Kim <hkim@llnl.gov>.

Published at the Differentiable Almost Everything Workshop of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. July 2023. Copyright 2023 by the author(s).

(Lahiri et al., 2023), but this method utilizes an enormous amount of memory (even though it is a sparse matrix) and is significantly inefficient because fetching the system matrix values from memory is much slower than computing these coefficient on the fly. In the case of cone-beam CT or high-resolution image reconstruction, the matrix size becomes excessively large, which makes the forward projection impractical. A neural network-based projector method (Gupta et al., 2018) is also less practical, due to the training requirement for each geometry. Most CT systems have a configurable geometry to optimize image quality and it is impractical to retrain with each configuration. The absence of a widely applicable, memory-efficient differentiable CT projection software tool remains a significant limitation.

This paper presents a new software library tool called LivermorE AI Projector (LEAP), providing differentiable CT forward projections to be easily integrated into neural network-based inference models. This tool provides both CPU- and GPU-based forward and back projections, utilizing a significantly small memory footprint, particularly with extensive GPU memory demands during neural network training and inference time. The main contributions are summarized as:

- The proposed tool provides differentiable projections to enable an end-to-end neural network training or inference pipeline for CT applications, without the need for significant memory footprint requirements.
- This supports forward and back projections for three widely used 3D scanner geometry types: parallel-beam, cone-beam, and a method to specify arbitrary locations and orientations of a set of source/detector pairs.
- This utilizes the highly accurate Separable Footprint (SF) projector model. Our implementation is also quantitatively accurate and all numerical values scale appropriately when changing the voxel sizes, detector sizes, etc.
- The proposed differentiable CT projections leverage the PyTorch interface and tensor formats, enabling seamless integration into existing neural network frameworks.
- The proposed tool is made available as open-source software, which can play a crucial role in driving innovation in various DL-based CT applications.
- The proposed tool can be used for implementing analytical or iterative reconstruction algorithms, which facilitates an integration of DL-driven reconstructions with conventional optimization-based methods.

2. Method

The proposed differentiable CT projectors are based on the core C++/CUDA module, Python binding and PyTorch interfaces, as shown in Figure 1. This section describes the projection models with the individual modules.

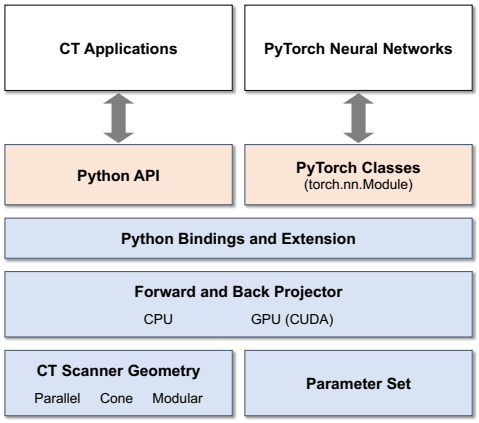

Figure 1: Overview of the proposed differentiable CT projector software architecture.

2.1. CT Projection Models and Geometries

The XCT forward model is given by the X-ray Transform which is a collection of line integrals through the unknown 3D volume (Natterer, 1986). The XCT model can be formulated as $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ where $\boldsymbol{x} \in \mathbb{R}^n$ is the volume of x-ray linear attenuation coefficients and $\boldsymbol{y}$ are the measured projection data. The forward model is denoted by $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ where the coefficients model the measurement process with the geometry of the CT system. XCT reconstruction is an inverse problem to recover the volume $\boldsymbol{x}$. Accurate reconstruction requires knowledge of the 3D position of every x-ray source and x-ray detector pixel pair which describes the geometry of a single measurement. Thus XCT forward models must be developed for the specific geometry of the imaging system. We will now discuss the CT geometries that our software package supports and provide some details on the algorithms we implemented.

We chose to implement three different 3D XCT geometry types in this software package: parallel-beam, axial cone-beam (planar or curved detector), and a flexible cone-beam geometry where uses may place x-ray sources and detectors at any position and orientation in space. These geometry types cover most XCT applications and future releases will include fan-beam and helical cone-beam geometries.

In practice, the reconstruction volume is composed of a 3D collection of voxels. Several mathematical models have been developed that model the XCT line integrals through a voxelized volume. The most popular and effective models are the Siddon method (Siddon, 1985), Joseph method (Joseph, 1982), Distance Driven (DD) method (Man & Basu, 2004), and the SF method (Long et al., 2010). The second two of these methods, DD and SF, model the finite width of the detector pixels and volume voxels, while the first two of these methods do not. Thus, although they are more computationally expensive to compute, the DD and SF methods

are more accurate and other methods have been shown to produce artifacts in some cases (Man & Basu, 2004). We chose to implement the Siddon and SF projector methods in our software package.

Any optimization-based method that relies on the calculation of gradients requires one to calculate the adjoint (or in the discrete case, the transpose) of the X-ray Transform (Barrett & Myers, 2013). The adjoint of the X-ray Transform is commonly referred to as *backprojection*. With some caveats (Zeng & Gullberg, 2000), one should employ methods for backprojection that are the exact transpose of the forward projection, referred to as *matched projectors*. Although there are some XCT reconstruction packages that utilize match projectors (Champley et al., 2022), most reconstruction packages (Aarle et al., 2015) violate this requirement because exact transposes are typically not as computationally efficient as other methods and if one stops the iterative reconstruction process early enough, artifacts will not appear. Since our goal here is to implement methods that are stable after over a thousand or more iterations, we chose to implement methods where the exact transpose is used. As an example of why the adjoint is needed, consider the following least squares $\widehat{x}_{LS} := \underset{x}{\operatorname{argmin}} \frac{1}{2}\|Ax - y\|^2$. Then the gradient of this cost function is given by $A^T(Ax - y)$.

All of our projectors are quantitatively accurate. The detector pixels and reconstruction voxels are specified in mm and the reconstruction volume units are in mm^{-1}.

Our parallel- and cone-beam implementations allow for a flexible specification of the geometry, including arbitrary 3D detector shifts and non-equispaced projection angles. If these still do not offer a flexible enough geometry, the user can specify the 3D location of every source-detector pair, as well as an arbitrary orientation of the detector.

Lastly, we also implemented forward/ back projector pairs for objects with cylindrical symmetry (Champley & Maddox, 2021). A special case of this is the Abel Transform which applies to parallel-beam geometries.

2.2. Core CUDA Implementation

Forward and back projection implementations are available for the three CT geometries described in the previous subsection using both the Siddon and SF projector models. These algorithms are implemented for the CPU using C/C++ and for NVIDIA GPUs using CUDA. Parallelization is done over the samples in the output space (CT projections for forward projection and reconstruction voxels for backprojection). Our CUDA implementation utilizes 3D threads and 3D texture memory is used for the input data (reconstruction volume for forward projection and CT projections for backprojection).

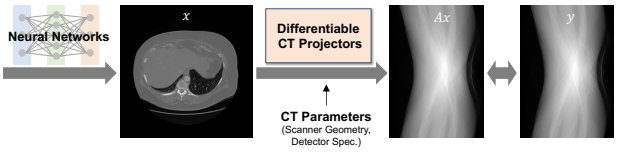

Figure 2: Model training and inference with our CT Projectors

CT projection data and the reconstruction volumes are both stored as contiguous 32-bit floating point arrays. Users simply provide the software with pointers to the relevant arrays. If the data are already on the GPU, then the appropriate method is implemented on this data, but if the data is not on the GPU (and the user want the calculation done on the GPU), the software copies the data to the GPU, performs the calculation, and then copies the result back to the CPU.

2.3. PyTorch Interfaces and CT Parameters

The forward and back projection in the core C++/CUDA module are provided as Python functions using the Python binding and the custom C++/CUDA extensions (Paszke et al., 2019). The main class is "Projector" which is derived from **torch.nn.Module** to enable automatic differentiation with PyTorch and the Python implementation.

The forward and back projections in the proposed library require a set of parameters specifying the CT scanner system. These CT parameters include scanner geometry type, number of projections, angular range or list of projection angles, detector pixel width and heights, source-to-detector distance, source-to-object distance, horizontal/vertical detector shift, number of voxels in each dimension, voxel sizes, and volume center position. They can be specified using set functions or a configuration file.

3. Sample Usage

Integration with Model Training and Inference A few recent DL methods addressing ill-posed CT reconstructions use forward projection models for data consistency (Zhou et al., 2021; Liu et al., 2022; Lahiri et al., 2023). These methods utilize a forward model as a loss function in the training or inference pipelines. While the detailed formulations differ from each other, the basic idea is to minimize the difference between the forward-projected data and original projection data, as illustrated in Figure 2. The loss function can be formulated as: $\arg\min_{\boldsymbol{x}\in\mathbb{R}^n} |\boldsymbol{Ax} - \boldsymbol{y}\|_2^2$.

Our proposed software library can be utilized as a differentiable forward model in existing DL frameworks. The provided PyTorch class enables seamless integration into existing **torch.nn.Module** classes. The Python code snippet below illustrates how to utilize our projection library during training or inference time:

```
import torch
```

```
from {our_projector_name} import Projector

proj = Projector(use_gpu=True, gpu_device="
    cuda:0", batch_size=1)
proj.load_param("parallel_beam.cfg")

for ind in range(0, n_iters):
  X, Y = data # retrieve data
  X_pred = model(X) # forward pass
  Y_pred = proj(X_pred) # forward project
  loss = loss_func(Y, Y_pred)
```

Listing 1: Example of our projector with a PyTorch model

End-to-End Reconstruction Pipeline While the proposed software library is primarily designed for forward modeling in neural network models, it can also be utilized for additional reconstruction processes or standalone reconstruction applications. The majority of the neural network inference-based methods in CT applications aims to improve ill-posed input images from conventional reconstruction algorithms (*e.g.*, Filtered Backprojection (FBP)) by removing artifacts and noise (Han & Ye, 2018; Zhou et al., 2021; Liu et al., 2022). In general, these initial reconstruction processes are performed using separate CT applications. Yet, the initial reconstruction process using our library has several advantages. First, it enables an integrated training pipeline with efficient memory usage. Second, this method enables to augment diverse, ill-posed input images given the training projection data. One can easily randomize the angular range (limited-angle CT) or the number of views (few-view CT) to generate diverse ill-posed input images.

Moreover, the predicted images from the inference models can be further improved by additional reconstruction algorithms (Kim et al., 2019). Our tool can also facilitate the seamless implementation of an end-to-end pipeline for the sinogram completion (Anirudh et al., 2018).

4. Experiments

Data Consistency with Inference Models We performed limited-angle CT experiments to demonstrate the benefit of the data consistency with the proposed differentiable projectors. We used a publicly available airport luggage dataset for automatic threat recognition (of Excellence at Northeastern University, 2014) that we split into 165 bags for training and the remaining 25 bags for test. The image dimension is 512^2 and the number of projections is 720 (parallel beam). To demonstrate the limited-angle CT, we randomly masked 120° out of 180° (60° available). We implemented a neural network model combining CT-Net (Anirudh et al., 2018) and U-Net (Han & Ye, 2018). Once the model was trained, we performed the sinogram completion and the iterative data consistency step using our projectors during the inference time. We compared the predicted images from the inference model and the final images after the sinogram completion and iterative refinement, with the ground truth images. The signal-to-noise-ratio score (PSNR) and structural similarity (SSIM) were used as the image quality metric. Figure 3 shows one of the predicted images directly from the model and its final image after the refinement. The refinement step with our projector led to an improvement in the averaged PSNR (dB) and SSIM from 35.486 and 0.905 to 36.350 and 0.911, respectively.

Figure 3: One of the predicted images from the inference model and its final image after the data consistency-based refinement with our projector. The PSNR/SSIM values are indicated at the bottom.

Geometry	Parallel		Cone	
Dimension	512^3/180	1024^3/720	512^3/180	1024^3/720
Ours	0.5/1.8 (1.5)	11.5/15.4 (8)	1.4/2.8 (1.5)	37.1/39.2 (11.1)
LTT	4.2 (-)	17.4 (-)	4.5 (-)	38.9 (-)

Table 1: Performance comparison (sec) between ours and LTT. (·) indicates the memory usage (GB). In our method, we report the times without and with the CPU-GPU data transfer. The dimension refers to the image dimension and the number of projections.

Performance and Memory Usage We report the forward projection time of our library and another publicly available reconstruction software tool (LTT) (Champley et al., 2022) on NVIDIA Tesla P100 with 16GB, as shown in Table 1. The angular ranges for parallel and cone beams are 180 and 360, respectively. Note that we did not list the memory usage for LTT because it is a user-specified parameter. The most GPU memory that LTT would use is enough to hold one copy of the projection data and volume data stored as 32-bit floats. This is also how much GPU memory LEAP requires as well. This small memory footprint requirement enables easy integration with existing neural network models on GPUs.

5. Conclusion

We presented an open-source software library providing differentiable XCT projectors. This library allows us to integrate forward modeling into DL frameworks as an end-to-end pipeline. This library enables fast forward and back projections with small memory footprint requirements.

Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-CONF-849839.

References

Aarle, W. V., Palenstijn, W. J., Beenhouwer, J. D., Altantzis, T., Bals, S., Batenburg, K. J., and Sijbers, J. The astra toolbox: a platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, 157:35–47, 2015.

Anirudh, R., Kim, H., Thiagarajan, J. J., Aditya Mohan, K., Champley, K., and Bremer, T. Lose the views: Limited angle ct reconstruction via implicit sinogram completion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6343–6352, 2018.

Anirudh, R., Kim, H., Thiagarajan, J. J., Mohan, A. K., and Champley, K. M. Improving limited angle ct reconstruction with a robust gan prior. *arXiv preprint arXiv:1910.01634*, 2019.

Barrett, H. H. and Myers, K. J. *Foundations of image science*. John Wiley & Sons, 2013.

Champley, K. and Maddox, B. Model based iterative reconstruction with the tilted abel transform. In *Digital Holography and Three-Dimensional Imaging*. Optica Publishing Group, 2021.

Champley, K. M., Willey, T. M., Kim, H., Bond, K., Glenn, S. M., Smith, J. A., Kallman, J. S., Brown, W. D., Seetho, I. M., Keene, L., Azevedo, S. G., McMichael, L. D., Overturf, G., and Martz, H. E. Livermore tomography tools: Accurate, fast, and flexible software for tomographic science. *NDT & E International*, 126: 102595, 2022. ISSN 0963-8695. doi: https://doi.org/10.1016/j.ndteint.2021.102595.

Gao, Q., Ding, R., Wang, L., Xue, B., and Duan, Y. Lrip-net: low-resolution image prior based network for limited-angle ct reconstruction. *IEEE Trans. Radiat. Plasma Med. Sci.*, 2022.

Gupta, H., Jin, K. H., Nguyen, H. Q., McCann, M. T., and Unser, M. CNN-based projected gradient descent for consistent ct image reconstruction. *IEEE Trans. Med. Imag.*, 37(6):1440–1453, Jun. 2018. ISSN 1558-254X. doi: 10.1109/TMI.2018.2832656.

H. E. Martz, J., Logan, C. M., Schneberk, D. J., and Shull, P. *X-ray Imaging: Fundamentals, Industrial Techniques and Applications*. CRC Press, 2016.

Han, Y. and Ye, J. C. Framing U-Net via deep convolutional framelets: Application to sparse-view CT. *IEEE Trans. Med. Imag.*, 37(6):1418–1429, 2018.

Hu, D., Zhang, Y., Liu, J., Luo, S., and Chen, Y. Dior: deep iterative optimization-based residual-learning for limited-angle ct reconstruction. *IEEE Trans. on Med. Imag.*, 41(7):1778–1790, 2022. doi: 10.1109/TMI.2022.3148110.

Joseph, P. An improved algorithm for reprojecting rays through pixel images. *IEEE Trans. Med. Imaging*, 1(3):192–196, 1982.

Kim, H., Anirudh, R., Mohan, A. K., and Champley, K. M. Extreme few-view ct reconstruction using deep inference. *Advances in Neural Information Processing Systems*, 33, 2019.

Lahiri, A., Maliakal, G., Klasky, M. L., Fessler, J. A., and Ravishankar, S. Sparse-view cone beam ct reconstruction using data-consistent supervised and adversarial learning from scarce training data. *IEEE Transactions on Computational Imaging*, 9: 13–28, 2023. doi: 10.1109/TCI.2022.3225680.

Liu, J., Anirudh, R., Thiagarajan, J. J., He, S., Mohan, K. A., Kamilov, U. S., and Kim, H. Dolce: A model-based probabilistic diffusion framework for limited-angle ct reconstruction, 2022.

Long, Y., Fessler, J. A., and Balter, J. M. 3d forward and back-projection for x-ray ct using separable footprints. *IEEE Transactions on Medical Imaging*, 29(11):1839–1850, 2010.

Man, B. D. and Basu, S. Distance-driven projection and backprojection in three dimensions. *Physics in Medicine & Biology*, 49 (11):2463, may 2004.

Natterer, F. *The Mathematics of Computerized Tomography*. John Wiley & Sons, Inc., Philadelphia, PA, USA, first edition, 1986.

of Excellence at Northeastern University, D. C. Alert to4 datasets for automated threat recognition. available at http://www.northeastern.edu/alert. 2014.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pp. 8024–8035. Curran Associates, Inc., 2019. URL https://pytorch.org/tutorials/advanced/cpp_extension.html.

Siddon, R. L. Fast calculation of the exact radiological path for a three-dimensional ct array. *Medical physics*, 12 2:252–5, 1985.

Zeng, G. L. and Gullberg, G. T. Unmatched projector/backprojector pairs in an iterative reconstruction algorithm. *IEEE Trans. Med. Imag.*, 19(5):548–555, 2000.

Zhang, Q., Hu, Z., Jiang, C., Zheng, H., Ge, Y., and Liang, D. Artifact removal using a hybrid-domain convolutional neural network for limited-angle computed tomography imaging. *Phys. Med. Biol.*, 65(15):155010, 2020.

Zhang, Y., Lv, T., Ge, R., Zhao, Q., Hu, D., Zhang, L., Liu, J., Zhang, Y., Liu, Q., Zhao, W., and Chen, Y. Cd-net: Comprehensive domain network with spectral complementary for dect sparse-view reconstruction. *IEEE Trans. Comput. Imag.*, 7: 436–447, 2021. doi: 10.1109/TCI.2021.3070184.

Zhou, B., Zhou, S. K., Duncan, J. S., and Liu, C. Limited view tomographic reconstruction using a cascaded residual dense spatial-channel attention network with projection data fidelity layer. In *IEEE Transactions on Medical Imaging*, 2021.

Zhou, K. C. and Horstmeyer, R. Diffraction tomography with a deep image prior. *Optics Express*, 28(9):12872–12896, 2020.