Precision Learning: Reconstruction Filter Kernel Discretization

C. Syben, B. Stimpel, K. Breininger, T. Würfl, R. Fahrig, A. Dörfler and A. Maier

Abstract—In this paper, we present substantial evidence that a deep neural network will intrinsically learn the appropriate way to discretize the ideal continuous reconstruction filter. Currently, the Ram-Lak filter or heuristic filters which impose different noise assumptions are used for filtered back-projection. All of these, however, inhibit a fully data-driven reconstruction neural network learning approach. In addition, the heuristic filters are not chosen in an optimal sense. To tackle this issue, we propose a formulation to directly learn the reconstruction filter. The filter is initialized with a Ramp filter as a strong pre-training and learned in frequency domain. We compare the learned filter with the Ram-Lak and the Ramp filter on a numerical phantom as well as on a real CT dataset. The results show that the network properly discretizes the continuous Ramp filter and converges towards the Ram-Lak solution. In our view these observations are interesting to gain a better understanding of deep learning techniques and traditional analytic techniques such as Wiener filtering and discretization theory. Furthermore, this will allow fully trainable data-driven reconstruction deep learning approaches.

Index Terms—Computed Tomography, Deep Learning, Filtered Back-Projection, Filter Discretization, Filter Learning

I. INTRODUCTION

Recently deep learning has shown promising results in the field of Computed Tomography (CT) reconstruction. In his perspective article, Wang [1] states that a reconstruction pipeline implemented as a deep neural network allows to access the capability of learning-based reconstruction. Wang identifies the data-driven knowledge-enhancing abilities as the strength of deep learning-based reconstruction. Würfl et al. [2] have proposed an implementation of the filtered back-projection algorithm (FBP) as a neural network. Similar to iterative reconstruction algorithms, their proposed implementation avoids explicitly storing the system matrix, which would render the algorithm infeasible. The approach utilized the data-driven capability by learning the compensation weights in case of limited-angle tomography. Hammernik et al. [3] proposed a two-level deep learning architecture to compensate for additional streak artifacts in the limited-angle tomography case. They showed that their approach allows for a joint optimization without any heuristic parameter tuning. In both
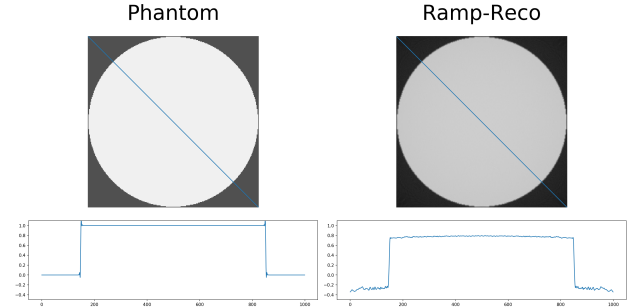
Fig. 1: Line profile through a 2D circle phantom. Ground truth (left) and FBP result based on Ramp filter (right).

approaches the necessary filtering to perform a FBP is done with a fixed layer using an analytical discretization.

The ideal filter for FBP can be derived using analytic reconstruction theory. Assuming an infinite number of projections and infinitely small detector pixels, it takes the form of the absolute value function in Fourier domain, commonly referred to as Ramp filter. In practice, however, Radon inversion has to be performed using a finite number of projections. This introduces discretization errors which need to be handled by replacing the Ramp filter with an appropriate discrete version. The occurring artifacts are commonly called cupping and dc shift artifacts [4], and an example of both is shown in Figure 1. The analytically derived discrete version of the Ramp filter is the well known Ram-Lak filter introduced by Ramachandran and Lakshminarayan [5]. Both, the ideal Ramp as well as the Ram-Lak filter are sensitive to noise. Thus, many different filters have been proposed which impose different noise assumptions. The most well known example is the Shepp-Logan filter [6], which incorporates a smoothing filter. However, while the choice of filter clearly matters, none of these heuristic filters are chosen in an optimal sense. CT vendors typically have specialized departments that aim at designing optimal filters for their users. The actual filter configurations are often deemed as company secrets. This shows the demand for specialized filters and raises the question whether data-driven methods can be used to deduce optimal filters.

Pelt et al. [7] proposed a method to learn such a data-dependent filter for the FPB algorithm. Their approach learns a filter approximation to increase reconstruction accuracy in specific cases. They optimize the filter with respect to the minimal error between the input data and the projection of the FBP reconstructed input. Pelt et al. use an efficient formulation and exponential binning to handle the size of the system

matanti. In order to do this, handling the discretization of the Ramp filter with respect to the details around the zero frequency as well as the boundaries is necessary.

We propose a formulation to learn a discrete optimal reconstruction filter directly in a deep learning context. We provide substantial evidence that a deep neural network will intrinsically learn the appropriate way to discretize the continuous filter. Furthermore, our proposed formulation leads to a straightforward implementation of the optimization using deep learning frameworks that automatically compute the gradient of the cost function using back-propagation.

II. METHOD & MATERIALS

First, we describe the filtered back-projection in continuous and discrete form for the parallel-beam geometry and introduce the theoretical filter kernel. Afterwards, the optimization problem and the related gradient to learn the filter kernel are derived. In the last Sections we describe our experiments and discuss them.

A. Filtered Back-Projection

The filtered back-projection (FBP) algorithm is an efficient solution to the reconstruction problem. First the projection data $p(s, \theta)$ are filtered by a convolution with the filter kernel

$$
\begin{aligned}
q(s, \theta) &= h(s) * p(s, \theta) \ , &(1) \\
h(s) &= \int |\omega| e^{2\pi\omega s} d\omega \ , &(2)
\end{aligned}
$$

subsequently, the filtered projection data $q(s, \theta)$ are back-projected to obtain the reconstruction result $f(x, y)$ with

$$
f(x, y) = \int_0^\pi q(s, \theta)|_{s=x \cdot \cos\theta + y \cdot \sin\theta} d\theta \ . \quad (3)
$$

In practice, we need a discrete description of the FBP presented in Eq. (1-3). The discrete reconstruction problem can be expressed by

$$
\boldsymbol{A}\boldsymbol{x} = \boldsymbol{p} \ , \quad (4)
$$

where $\boldsymbol{A} \in \mathbb{R}^{N \times M \cdot P}$ is the system matrix, $\boldsymbol{x} \in \mathbb{R}^N$ is the volume and $\boldsymbol{p} \in \mathbb{R}^{M \cdot P}$ are the projections. Since the discrete representation of the Radon transform $\boldsymbol{A}$ is a tall matrix, it has no inverse. This means every reconstruction formula is, therefore, a unique pseudo inverse of $\boldsymbol{A}$,

$$
\boldsymbol{x} = \underbrace{\boldsymbol{A}^\top}_{\text{Back-projection (Eq. 3)}} \underbrace{(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}}_{\text{Filter (Eq.1)}} \boldsymbol{p} \ . \quad (5)
$$

B. Learning the Filter Kernel

From the convolution theorem it is clear that a convolution in spatial domain is equal to a multiplication in the frequency space. Therefore, we can reformulate the filter $(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}$ in form of a diagonal matrix in Fourier domain

$$
\boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{F}^H \boldsymbol{K} \boldsymbol{F} \boldsymbol{p} \ , \quad (6)
$$

where $\boldsymbol{F} \in \mathbb{R}^{N \times N}$ and $\boldsymbol{F}^H \in \mathbb{R}^{N \times N}$ represents the Fourier and inverse Fourier transform, respectively. $\boldsymbol{K} \in \mathbb{R}^{N \times N}$ is a diagonal matrix representing the filter in frequency domain. To learn the filter matrix $\boldsymbol{K}$, which is a discrete approximation of

the Ramp filter, we use Eq. 6 to formulate a objective function $f(\boldsymbol{K})$ as a least-square minimization problem:

$$
f(\boldsymbol{K}) = \frac{1}{2}\|\boldsymbol{A}^\top \boldsymbol{F}^H \boldsymbol{K} \boldsymbol{F} \boldsymbol{p} - \boldsymbol{x}\|_2^2 \ . \quad (7)
$$

The gradient to our objective function $f(\boldsymbol{K})$ in Eq. 7 with respect to $\boldsymbol{K}$ is

$$
\frac{\partial f(\boldsymbol{K})}{\partial \boldsymbol{K}} = \boldsymbol{F}\boldsymbol{A} \underbrace{\underbrace{(\boldsymbol{A}^T \boldsymbol{F}^H \boldsymbol{K} \boldsymbol{F} \boldsymbol{p} - \boldsymbol{x})}_{\text{Error}} \underbrace{(\boldsymbol{F}\boldsymbol{p})^\top}_{l-1}}_{\text{Back-propagation}} \ . \quad (8)
$$

Note that this analytical gradient also has an interpretation in a neural network learning context. To describe Eq. 8 using the terms of back-propagation: Eq. 6 can be regarded as a network with input $\boldsymbol{p}$ and layers $\boldsymbol{F}$, $\boldsymbol{K}$, $\boldsymbol{F}^H$ and $\boldsymbol{A}^T$ with the identity as activation function between layers. $\boldsymbol{K}$ is the only layer containing trainable weights. Then, $\hat{\boldsymbol{x}} = \boldsymbol{A}^T \boldsymbol{F}^H \boldsymbol{K} \boldsymbol{F} \boldsymbol{p}$ is the forward pass through the network following the considerations by Würfl et al. [2]. The gradient of the error function (Eq. 7) with respect to $\boldsymbol{K}$ is computed by multiplying two factors: 1) the partial derivative of the error function with respect to the output of the layer and 2) the transpose of the output of the previous layer $(l-1)$. The output of $l-1$ is readily described by $\boldsymbol{F}\boldsymbol{p}$. Using the recursive formulation of back-propagation, we yield $\boldsymbol{F}\boldsymbol{A}(\hat{\boldsymbol{x}} - \boldsymbol{x})$, with $(\hat{\boldsymbol{x}} - \boldsymbol{x})$ being the derivative of the error function. We consider these observations as interesting, as this gradient would be computed automatically in a deep learning framework such as TensorFlow.

C. Experiments

We implemented the cost function and the analytically derived gradient in CONRAD [8] and used stochastic gradient descent to learn the filter. Unmatched projectors were used for the reconstruction (pixel driven) and the forward projections (ray driven). 0 The filter matrix $\boldsymbol{K}$ is initialized with a slightly modified frequency domain representation of the ideal Ramp filter. By doubling the width of the zero-valued part of the ramp we emphasize the cupping artifacts. This serves exclusively to show the learning capability. For the training we use 10 numerical disc phantoms with increasing radii rendered on a 512×512 pixel grid. The learned filter is evaluated on a 512×512 slice of a real CT dataset showing the head of a pig acquired at Stanford University. For evaluation we compare the filtered back-projection using the modified Ramp Filter, the Ram-Lak and the learned filter, in the following referred to as Ramp-reco, Ram-Lak-reco and Learned-reco, respectively. The quantitative evaluation is done on the absolute difference between the ground truth (GT) and the respective reconstruction result. These difference images are evaluated using the mean, minimum and maximum difference as well as the standard deviation (std. dev.). Note that we use the original reconstruction as the ground truth for the pig experiment.

III. RESULTS

A. Qualitative Results

Fig. 2 shows the results of the FBP on a numerical disc phantom, which was used for the training. Additionally, the
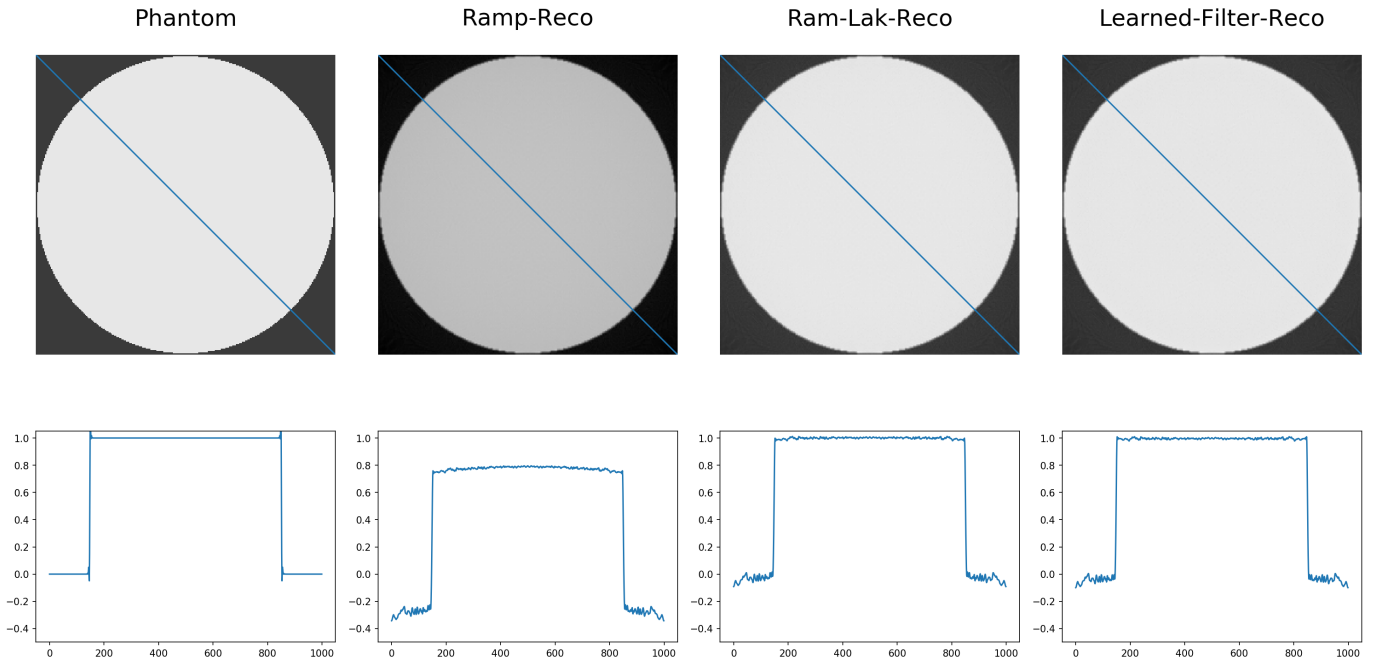
Fig. 2: Reconstructions and respective line profile plots of numeric circle phantom using different filters. All images have the same window/level.
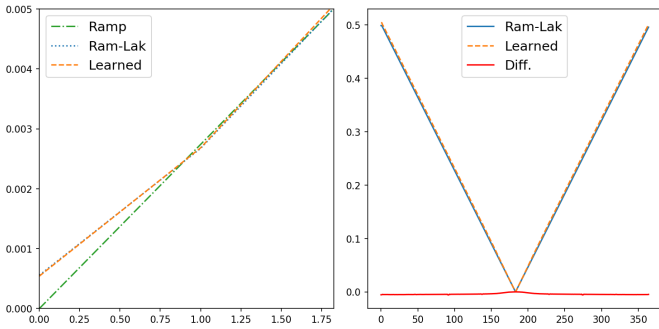

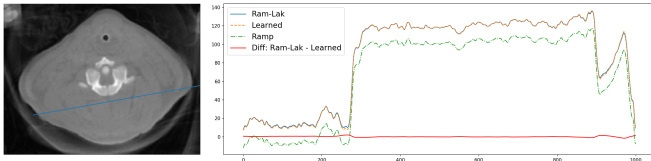Fig. 3: Zero frequency behavior of the Ramp, Ram-Lak and learned filter.

Fig. 4: Reconstruction of a pig dataset filtered with Ramp, Ram-Lak and learned filter. The difference between the Ram-Lak and the learned filter is plotted in red.

results of Ramp-Reco , the Ram-Lak-Reco and the Learned-Filter-Reco and their respective line profiles are presented. The Ramp-Reco leads to cupping artifacts in the homogeneous area of the disc that can be observed in Fig. 2 as well as in the respective line profile. The line profile for the Ram-Lak-Reco illustrates that the homogeneous area of the disc is reconstructed properly. The Learned-Filter-Reco reconstructs the homogeneous area properly as well. Comparing the line

profiles of the Ram-Lak-Reco with the line profile from the Learned-Reco, the Ram-Lak-Reco shows nearly a straight line while the line profile of the Learned-Reco still shows small deviations from the ideal line. Fig.3 show the differences of the Ramp, Ram-Lak and learned filter around the zero frequency and on the full spectrum. In Fig. 4 the reference reconstruction is shown and the respective line-profiles from the different reconstructions using the Ramp, Ram-Lak and learned Filter are plotted. Both line profiles, aligned over the same homogeneous area of the pig, show a similar behavior and no cupping.

B. Quantitative Results

In Tab. I the measurements of the absolute difference between the GT and the respective reconstructions are shown. The absolute mean error of the Ramp-Reco is at 23.5%, while the absolute mean error of the Learned-Filter-Reco is at 2.3%. The lowest absolute mean error can be observed by the Ram-Lak-Reco. The standard deviation as well as the maximum value of the Ramp-Reco are higher compared to the other two reconstructions. The Ram-Lak-Reco and the Learned-Filter-Reco show a similar standard deviation of the absolute difference as well as the absolute maximum error. In Tab. II the absolute error measurements between the GT and the respective reconstructions are presented. Both the Ram-Lak-Reco as well as the Learned-Filter-Reco exhibit a similar absolute mean error. Also the standard deviation of the absolute error as well as the absolute maximum error are closely together.

IV. DISCUSSION

The evaluation of the numerical disc phantom shows that homogeneous areas can be reconstructed properly without cup-

TABLE I: Measurements of the absolute difference between GT and respective circle reconstruction (in percent).

	mean	std. dev.	min	max
Ramp-reco	0.235	0.07	0.001	0.596
Ram-Lak-reco	0.01	0.031	0	0.41
Learned-reco	0.023	0.03	6.76E-09	0.409

TABLE II: Measurements of the absolute difference between GT and respective pig reconstruction in Hounsfield units (HU).

	mean	std. dev.	min	max
Ram-Lak-reco	66.99	61.401	6.10E-5	1634.82
Learned-reco	83.53	68.06	8.39E-5	1685.70

ping artifacts when using the learned filter. However, the line profile reveals that the result is not as good as with the Ram-Lak filter. A possible explanation for this difference is our way of implementing the optimization directly and configuring the stochastic gradient parameters heuristically. Also note that our current training is performed with only 20 epochs and our data set consists only of 10 cylinders of different diameter. This is only a coarse approximation of the ideal training set. We assume that using a basis that spans the entire domain of x will do much better for this job. This will be subject of future experiments. Still, we consider 10 training samples as a good start for estimating such a complex relation that generalizes to other much more complex objects such as the pig data set. The minor difference between the Learned-reco and the Ram-Lak-reco in the real CT data experiment prove that the learned filter is not object dependent and not over-fitted to our training data. Utilizing discs with varying radii introduces two properties to the training process. First, narrow discs model the Dirac-impulse. Secondly, discs with larger radii exhibit large homogeneous regions. Occurring cupping and dc-shift artifacts in the homogeneous area will lead to strong gradients, which appear due to the wrong discretization. As a consequence, the weights converge towards the Ram-Lak solution, which is nothing else than learning the proper discretization of the continuous Ramp. Using the ideal Ramp less cupping is observed. This renders the dc-shift as the dominant artifact to compensate for. We expect that augmentation approaches of this method will lead to filters that are invariant / less prone to noise characteristics imposed by CT physics intrinsically. Thus, we believe that noise augmentation will lead to filters similar to the Shepp-Logan filter. Augmentation in this context will have a very similar result as a Wiener filter that is optimal given certain noise properties. Different approaches to learn the reconstruction filter were published in the past. Floyd [9] successfully learned the discrete version of the ramp filter for SPECT image reconstruction. However, compared to our presented method, their approach learns the filter in spatial domain using a neural network with fully connected layers. As a consequence of the system design, Floyd reported practical issues implementing the huge amount of trainable weights at the time of the publication. Furthermore, no detailed analysis of the discretization properties of the learned filter was performed. Even though similar approaches to learn the reconstruction filter were published, e.g. [7], [9], none of

them explicitly uses domain knowledge to design the network topology. In contrast, our derivation of the network topology is based on the continuous analytical problem description. Furthermore, the transition to the discrete filter is intrinsically solved due to the discrete nature of the neural networks. In our view these observations are interesting to gain a better understanding of deep learning techniques and traditional analytic techniques such as Wiener filtering and discretization theory. To the best of our knowledge, we did not observe such links between analytical signal processing theory and deep learning so far.

V. Conclusion

We presented an approach to learn the discrete optimal reconstruction filter directly from the continuous Ramp filter. We have shown that the learning approach will automatically compensate for the errors inflicted by the discretization in an L2-sense optimal way with respect to our given training data. This is achieved by formulating a cost function to learn the filter in the frequency domain. This enables us to initialize the filter with the ideal Ramp, which can be seen as a very strong pre-training. Furthermore, the formulation can be straightforward transfered to a neural network architecture. Combining the proposed solution with the deep neural network suggested by Würfl et al. enables us to provide a fully trainable data-driven reconstruction deep learning approach. In future work, we want to apply noise models to the training data to learn an optimal discrete filter which is less sensitive to noise.

References

[1] Ge Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.

[2] Tobias Würfl, Florin Cristian Ghesu, Vincent Christlein, and Andreas Maier, "Deep Learning Computed Tomography," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Proceedings, Part III*, 2016, vol. 3, pp. 432–440.

[3] Kerstin Hammernik, Tobias Würfl, Thomas Pock, and Andreas Maier, "A deep learning architecture for limited-angle computed tomography reconstruction," in *Bildverarbeitung für die Medizin 2017: Algorithmen - Systeme - Anwendungen. Proceedings des Workshops vom 12. bis 14. März 2017 in Heidelberg*, Berlin, Heidelberg, 2017, pp. 92–97, Springer Berlin Heidelberg.

[4] Avinash C. Kak and Malcolm Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.

[5] GN Ramachandran and AV Lakshminarayanan, "Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of fourier transforms," *Proceedings of the National Academy of Sciences*, vol. 68, no. 9, pp. 2236–2240, 1971.

[6] Lawrence A Shepp and Benjamin F Logan, "The fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 21–43, 1974.

[7] Daniël M Pelt and Kees Joost Batenburg, "Improving filtered backprojection reconstruction by data-dependent filtering," *IEEE Transactions on Image Processing*, vol. 23, no. 11, pp. 4750–4762, 2014.

[8] Andreas Maier, Hannes Hofmann, Martin Berger, Peter Fischer, Chris Schwemmer, Haibo Wu, Kerstin Müller, Joachim Hornegger, Jang-Hwan Choi, Christian Riess, Andreas Keil, and Rebecca Fahrig, "CONRAD - A software framework for cone-beam imaging in radiology," *Medical Physics*, vol. 40, no. 11, 2013.

[9] C. E. Floyd, "An artificial neural network for spect image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 10, no. 3, pp. 485–487, Sep 1991.