NEWS & VIEWS   Open Access

Deep learning enables nanoscale X-ray 3D imaging with limited data

Chonghang Zhao [1] and Hanfei Yan [1]✉

Abstract
Deep neural network can greatly improve tomography reconstruction with limited data. A recent effort of combining ptycho-tomography model with the 3D U-net demonstrated a significant reduction in both the number of projections and computation time, and showed its potential for integrated circuit imaging that requires high-resolution and fast measurement speed.

X-ray tomography is a non-destructive imaging technique that provides three-dimensional (3D) structural information about an object. It has many applications in various scientific fields. The technique involves taking a series of projection images as the sample rotates, and then using a mathematical algorithm to compute the volumetric reconstruction. The quality of the reconstruction depends on both the data and the computation method. To achieve a given resolution, the number of projections must meet the Crowther criterion, which means that hundreds to thousands of projections are required under optimal conditions[1]. Regarding tomography algorithms, they can be mainly divided into two categories: direct methods that utilize back-projection or the connection between Radon and Fourier transforms[2–4], and model-based methods that iteratively solve an optimization problem[5,6]. Direct methods require less computation but reconstructions are more susceptible to artifacts with poor-quality data. In contrast, iterative methods are more robust but are less efficient.

In some cases, it is not possible to meet the data requirements due to concerns about radiation dose or geometrical constraints. For example, the tomography of integrated circuits (ICs) is limited by its plate-like geometry, which restricts the collection of projection images to a limited angular range. This leads to a well-known missing edge problem in tomography. In addition, it is

desirable to take as few projections as possible to speed up the measurement and cover a large area. Achieving high-quality reconstruction with fast measurement and light computation has been a focus of research for a long time.

The recent emergence of deep neural networks (DNNs) has opened up new opportunities to tackle this challenge. DNNs enable machines to learn complex and implicit knowledge, allowing them to provide artificial intelligence (AI). By harnessing the power of DNNs, researchers have been able to achieve nearly perfect reconstruction with incomplete datasets, a task that was previously difficult or impossible with conventional methods[7–9].

DNNs can be used in many ways in tomography. They can act as a black box to replace the iterative solver and directly map the measured data to the target image[9–12]. They can also serve as a post-image processing tool to improve the image after the reconstruction[7,8]. Additionally, they can perform sinogram inpainting to fill in missing information in the measurement or regularize a solution in a model-based approach to preserve desired properties[13–15].

A recent paper by Z. Wu et al. demonstrates another successful marriage of DNNs and tomography[16]. The team from MIT and Argonne National Laboratory successfully reconstructed an IC sample using only 21 projections in a range of 140 degrees. Their proposed method, named RAPID, greatly reduces the required amount of data and speeds up computation time by 140 times while retaining fine details of the object and achieving a similar quality to that obtained from a full dataset with 349 projections.
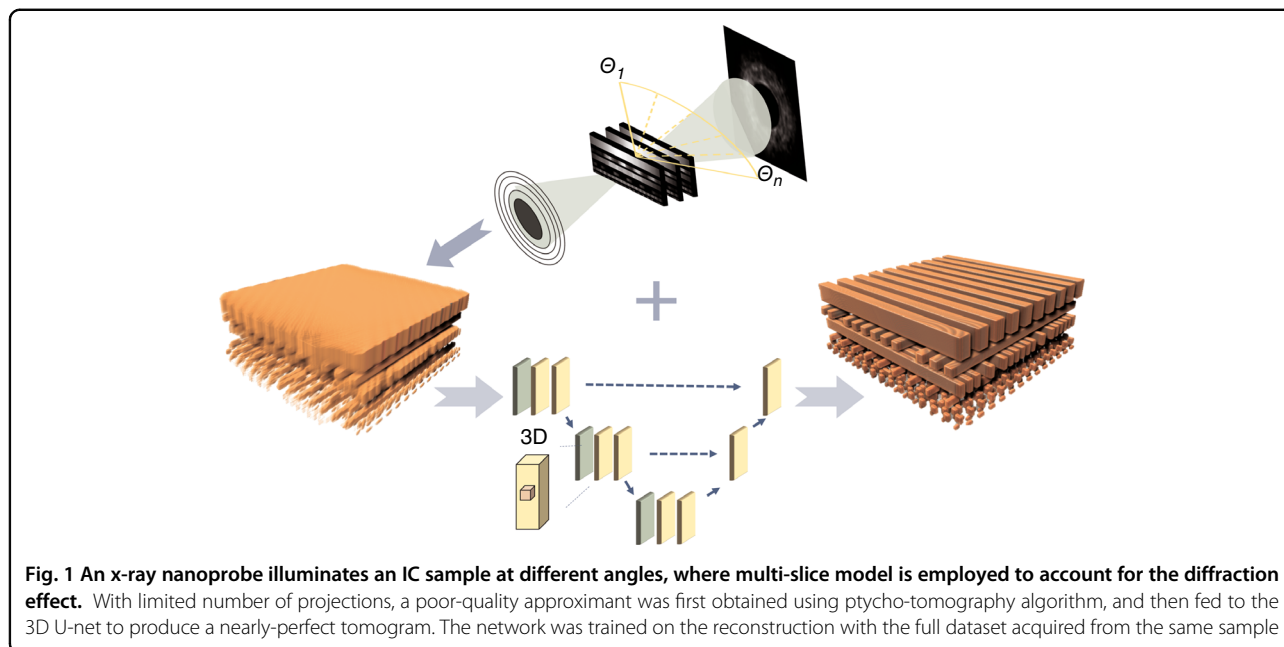
Correspondence: Hanfei Yan (hyan@bnl.gov)
[1]National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA

Fig. 1 An x-ray nanoprobe illuminates an IC sample at different angles, where multi-slice model is employed to account for the diffraction effect. With limited number of projections, a poor-quality approximant was first obtained using ptycho-tomography algorithm, and then fed to the 3D U-net to produce a nearly-perfect tomogram. The network was trained on the reconstruction with the full dataset acquired from the same sample

RAPID differs from existing efforts in two aspects (See Fig. 1). First, it employs a multi-slice propagation method to model the measurement, which accounts for the diffraction effect inside the object[17,18]. This potentially removes the limitation on object thickness imposed by the depth of view and allows for the achievement of spatial resolution better than that of the optical system. Using a Fresnel Zone plate with an outermost zone width of 50 nm, the team was able to achieve a voxel resolution of 14 nm.

Second, RAPID implements a 3D U-net with atrous convolutions for larger receptive fields, allowing the network to capture 3D features more efficiently. The combination of a 3D physical model and a 3D network makes it possible to dramatically reduce both the number of projections and computation time for a tomogram. It also eliminates the normalization issue between slices sometimes seen with 2D methods.

In RAPID, supervised learning is used. The dataset acquired from the IC sample was split into two parts: one for training and one for testing. A high-resolution reconstruction was generated using the conventional two-step approach of first reconstructing the 2D projection images and then performing tomographic reconstruction. This reconstruction was used as the ground truth for training. Various metrics were used to evaluate the performance of the method, and RAPID consistently outperformed conventional filtered back-propagation (FBP) and simultaneous algebraic reconstruction techniques (SART).

The proposed strategy is to train the network on a subset of the sample where sufficient measurements have been conducted and good results have been achieved with conventional methods. The network can then be applied to the rest of the sample where a very sparse dataset has been collected to significantly speed up the process. One drawback is that new training may be required for a different sample, but transfer learning may reduce the effort if the features are similar.

With the power of DNNs, which enforce solutions to meet expectations based on prior knowledge, visually appealing results can be obtained from incomplete or sometimes "trash" datasets. This is the case with RAPID and other AI-enabled algorithms. These methods can greatly reduce the effort required to measure similar samples while still maintaining quality. However, because the network is trained on prior knowledge, the solution is biased toward that by design. One caveat is that if there is a new feature that was not seen during training, the network may ignore or misinterpret it. This is not an issue if we do not expect any surprises from a sample, such as in a screening application. However, in scientific research where unknown features are of interest, this raises several open questions.

Is the solution unique? Could there be another visually appealing solution that is equally probable? Can uncertainty be assigned to different locations of the object to indicate which portions are more "guessed" by AI? Unlike conventional methods where a "bad" result is easily recognizable, an AI-enabled tool can generate an output that always looks "good". To what degree can we trust it? Appropriate metrics may need to be defined to describe the confidence level.

References

1. Crowther, R. A., DeRosier, D. J. & Klug, A. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proc. R. Soc. A Math., Phys. Eng. Sci.* **317**, 319–340 (1970).
2. Radon, J. On the determination of functions from their integral values along certain manifolds. *IEEE Trans. Med. Imaging* **5**, 170–176, https://doi.org/10.1109/tmi.1986.4307775 (1986).
3. Dowd, B. A. et al. Developments in synchrotron x-ray computed micro-tomography at the national synchrotron light source. Proceedings of the SPIE 3772, Conference on Developments in X-Ray Tomography II. Denver: SPIE, 1999, 224–236.
4. Kak, A. C. & Slaney, M. Principles of Computerized Tomographic Imaging. (Society for Industrial and Applied Mathematics, Philadelphia, 2001).
5. Andersen, A. H. & Kak, A. C. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the art algorithm. *Ultrasonic Imaging* **6**, 81–94, https://doi.org/10.1016/0161-7346(84)90008-7 (1984).
6. Nuyts, J., Michel, C. & Dupont, P. Maximum-likelihood expectation-maximization reconstruction of sinograms with arbitrary noise distribution using NEC-transformations. *IEEE Trans. Med. Imaging* **20**, 365–375, https://doi.org/10.1109/42.925290 (2001).
7. Zhang, Z. C. et al. A sparse-view CT reconstruction method based on combination of densenet and deconvolution. *IEEE Trans. Med. Imaging* **37**, 1407–1417, https://doi.org/10.1109/tmi.2018.2823338 (2018).
8. Jin, K. H. et al. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522, https://doi.org/10.1109/tip.2017.2713099 (2017).
9. Zhu, B. et al. Image reconstruction by domain-transform manifold learning. *Nature* **555**, 487–492, https://doi.org/10.1038/nature25988 (2018).
10. Yang, X. G. et al. Tomographic reconstruction with a generative adversarial network. *J. Synchrotron Radiat.* **27**, 486–493, https://doi.org/10.1107/s1600577520000831 (2020).
11. Tan, C. et al. Image reconstruction based on convolutional neural network for electrical resistance tomography. *IEEE Sens. J.* **19**, 196–204, https://doi.org/10.1109/jsen.2018.2876411 (2019).
12. McCann, M. T., Jin, K. H. & Unser, M. Convolutional neural networks for inverse problems in imaging: a review. *IEEE Signal Process. Mag.* **34**, 85–95, https://doi.org/10.1109/msp.2017.2739299 (2017).
13. Lee, H. et al. Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Trans. Radiat. Plasma Med. Sci.* **3**, 109–119, https://doi.org/10.1109/trpms.2018.2867611 (2019).
14. Wu, D. F. et al. Iterative low-dose CT reconstruction with priors trained by artificial neural network. *IEEE Trans. Med. Imaging* **36**, 2479–2486, https://doi.org/10.1109/tmi.2017.2753138 (2017).
15. Xiang, J. X., Dong, Y. G. & Yang, Y. J. FISTA-Net: learning a fast iterative shrinkage thresholding network for inverse problems in imaging. *IEEE Trans. Med. Imaging* **40**, 1329–1339, https://doi.org/10.1109/tmi.2021.3054167 (2021).
16. Wu, Z. L. et al. Three-dimensional nanoscale reduced-angle ptycho-tomographic imaging with deep learning (RAPID). *eLight* **3**, 7, https://doi.org/10.1186/s43593-022-00037-9 (2023).
17. Maiden, A. M., Humphry, M. J. & Rodenburg, J. M. Ptychographic transmission microscopy in three dimensions using a multi-slice approach. *J. Optical Soc. Am. A* **29**, 1606–1614, https://doi.org/10.1364/josaa.29.001606 (2012).
18. Li, P. & Maiden, A. Multi-slice ptychographic tomography. *Sci. Rep.* **8**, 2049, https://doi.org/10.1038/s41598-018-20530-x (2018).