

Towards full-stack deep learning-empowered data processing pipeline for synchrotron tomography experiments

Zhen Zhang,^{1,3} Chun Li,^{2,3} Wenhui Wang,² Zheng Dong,^{2,*} Gongfa Liu,¹ Yuhui Dong,² and Yi Zhang^{2,*}

*Correspondence: dongz@ihep.ac.cn (Z.D.); zhangyi88@ihep.ac.cn (Y.Z.)

Received: June 16, 2023; Accepted: November 13, 2023; Published Online: November 16, 2023; <https://doi.org/10.1016/j.xinn.2023.100539>

© 2023 The Author(s). This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Tomography experiments at future synchrotron beamlines face exascale big data challenges.
- Deep learning is employed for various data processing tasks at current synchrotron tomography beamlines.
- Future comprehensive and elongated data analysis processes require a full-stack deep learning pipeline.
- The data-driven full-stack deep learning pipeline based on an intelligent scheduling center (ISC) holds the key to solve the big data challenges.



Towards full-stack deep learning-empowered data processing pipeline for synchrotron tomography experiments

Zhen Zhang,^{1,3} Chun Li,^{2,3} Wenhui Wang,² Zheng Dong,^{2,*} Gongfa Liu,¹ Yuhui Dong,² and Yi Zhang^{2,*}

¹National Synchrotron Radiation Laboratory, University of Science and Technology of China, Hefei 230029, China

²Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

³These authors contributed equally

*Correspondence: dongz@ihep.ac.cn (Z.D.); zhangyi88@ihep.ac.cn (Y.Z.)

Received: June 16, 2023; Accepted: November 13, 2023; Published Online: November 16, 2023; <https://doi.org/10.1016/j.xinn.2023.100539>

© 2023 The Author(s). This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Citation: Zhang Z., Li C., Wang W., et al., (2024). Towards full-stack deep learning-empowered data processing pipeline for synchrotron tomography experiments. *The Innovation* 5(1), 100539.

Synchrotron tomography experiments are transitioning into multifunctional, cross-scale, and dynamic characterizations, enabled by new-generation synchrotron light sources and fast developments in beamline instrumentation. However, with the spatial and temporal resolving power entering a new era, this transition generates vast amounts of data, which imposes a significant burden on the data processing end. Today, as a highly accurate and efficient data processing method, deep learning shows great potential to address the big data challenge being encountered at future synchrotron beamlines. In this review, we discuss recent advances employing deep learning at different stages of the synchrotron tomography data processing pipeline. We also highlight how applications in other data-intensive fields, such as medical imaging and electron tomography, can be migrated to synchrotron tomography. Finally, we provide our thoughts on possible challenges and opportunities as well as the outlook, envisioning selected deep learning methods, curated big models, and customized learning strategies, all through an intelligent scheduling solution.

INTRODUCTION

The methodology of tomography is a three-dimensional (3D) characterization technique capable of revealing internal structures and functions of matter using various probe types, such as X-rays, electrons, neutrons, etc. Tomography has been widely adopted in various scientific fields since its emergence in the early 20th century. In biology,^{1–7} tomography unveils 3D assembly and interaction mechanisms from macromolecules, organelles and cells to tissues and living organisms, making it a powerful tool for uncovering the mysteries of life. In medical science,^{8–12} tomography facilitates internal examination of lesion areas, aiding in the subsequent diagnosis and treatment process. In materials science,^{13–18} tomography, when combined with *in situ* methods, is used to analyze structures, identify defects, and uncover the dynamics within materials, greatly benefitting the discovery and study of new-generation functional materials.

The boundary of tomography is pushed further with the commissioning of large scientific research facilities, such as new-generation synchrotron light sources,^{19–21} which greatly improves the ability to characterize matter at extreme spatial and temporal resolution. This improvement is mostly due to the ultra-bright, time-resolved, and highly coherent nature of X-ray sources generated in the beamlines of new-generation synchrotron facilities, which enables synchrotron tomography. Advanced imaging techniques, such as ptychography²² and holography,²³ are being developed to further promote the spatial resolution to below 10 nm. At the ID16A²⁴ beamline in European Synchrotron Radiation Facility (ESRF), the spatial resolving capability of X-ray nano-holography experiments is reaching cryoelectron microscopy (cryo-EM) levels²⁴ with an incomparable time resolving advantage for *in situ* characterization of large millimeter-size samples. To image even larger samples, the hierarchical phase-contrast tomography (HiP-CT)²⁵ technique offered by the BM05 beamline at ESRF has demonstrated the capability to image the entire lung of a coronavirus disease 2019 (COVID-19) patient with 1- μm resolution, which unlocks the resolving power on subcellular features. In modern synchrotron tomography, there is a rising interest in multimodal characterizations, which facilitates a better understanding of structural and functional relationships. For example, at nanoprobe beamlines such as Hard X-ray Nanoprobe (HXN; Brookhaven National Laboratory)²⁶ and P06 (Deutsches Elektronen-Synchrotron),²⁷ nanoscale X-ray spectroscopy (fluorescence, absorption near-edge structure), diffraction, ptychography, and tomography experiments can be performed simultaneously to acquire both chemical and

structural information; small-angle X-ray scattering tensor tomography and wide-angle X-ray diffraction tensor tomography are often combined with high-resolution absorption tomography experiments to examine the orientation and strain information in textured samples. Another important trend is dynamic tomography, where *in situ* or dynamic experiments are being undertaken as routine in many beamlines. For example, the TomoCAT beamline (Paul Scherrer Institute)²⁸ is capable of providing serial tomography characterization techniques at a rate of 20 Hz.

These changes in synchrotron tomography inevitably lead to explosive growth in data volume and dimension as well as elongated processing pipelines, imposing a tremendous burden on the data processing end.²⁹ Some beamlines are expected to generate petabytes (PBs) of data per day, becoming some of the largest data sources for scientific experiments. The demand on real-time analysis will be more imperative than ever. Therefore, it is essential to address the big data challenge, optimize the tomographic data processing pipeline, and ultimately improve experimental characterization capabilities because conventional techniques are still struggling to cope with such workloads.

Today, scientific research is entering the “fourth paradigm,”³⁰ powered by numerous data-driven approaches. Among them, deep learning (DL) techniques have been widely acknowledged as a highly accurate and efficient data processing method compared with conventional techniques. Over the last few decades, DL-based network models have achieved outstanding performance on various downstream tasks, such as classification, object recognition, instance segmentation, image generation, and more. From the emergence of artificial neural networks (ANN) based on multilayer perceptron (MLP) architecture in the 1990s to deep neural networks constructed by convolutional neural network (CNN) and today's transformer architecture that showcases superior feature extraction capabilities, DL architectures with higher efficiency and accuracy have proliferated. One of the advantages of DL is the ability to use massive amounts of data for learning, which maps the principles between inputs and outputs while overcoming problems caused by formulaic solutions. Therefore, researchers are exploring the possibilities of applying the “scientific big data + DL” paradigm to synchrotron tomography. Integrating DL into the synchrotron tomography data processing pipeline is believed to be the ultimate solution to tackle the challenges encountered in high-resolution, multimodal, cross-scale, and dynamic experiments.

Taking 3D reconstruction as a pivotal module, DL applications on the synchrotron tomography data processing pipeline can be categorized as three workloads in succession (Figure 1): image processing before reconstruction (preprocessing), 3D reconstruction optimization, and scientific application-oriented data processing on reconstructed data. The preprocessing workload usually involves image stitching and a series of correction tasks to tackle the misalignment of the projection images due to system vibrations and offsets. The optimization process usually concentrates on defect mitigation, specifically on optimization of the 3D reconstruction process for sparse-view datasets, as well as denoising on reconstructed volumes. Finally, scientific application-oriented data processing on reconstructed data task directly relates to specific scientific goals.

This review aims to provide a comprehensive overview of DL applications in the synchrotron tomography data processing pipeline (STDPP). It will present and discuss each element of the pipeline, providing guidance on tomographic data processing for both inexperienced and skilled researchers in the data-driven era. In addition, other data-intensive research fields, such as medical imaging and electron tomography (ET), are similar to synchrotron tomography in terms of application scenarios. With DL being more thoroughly studied in these

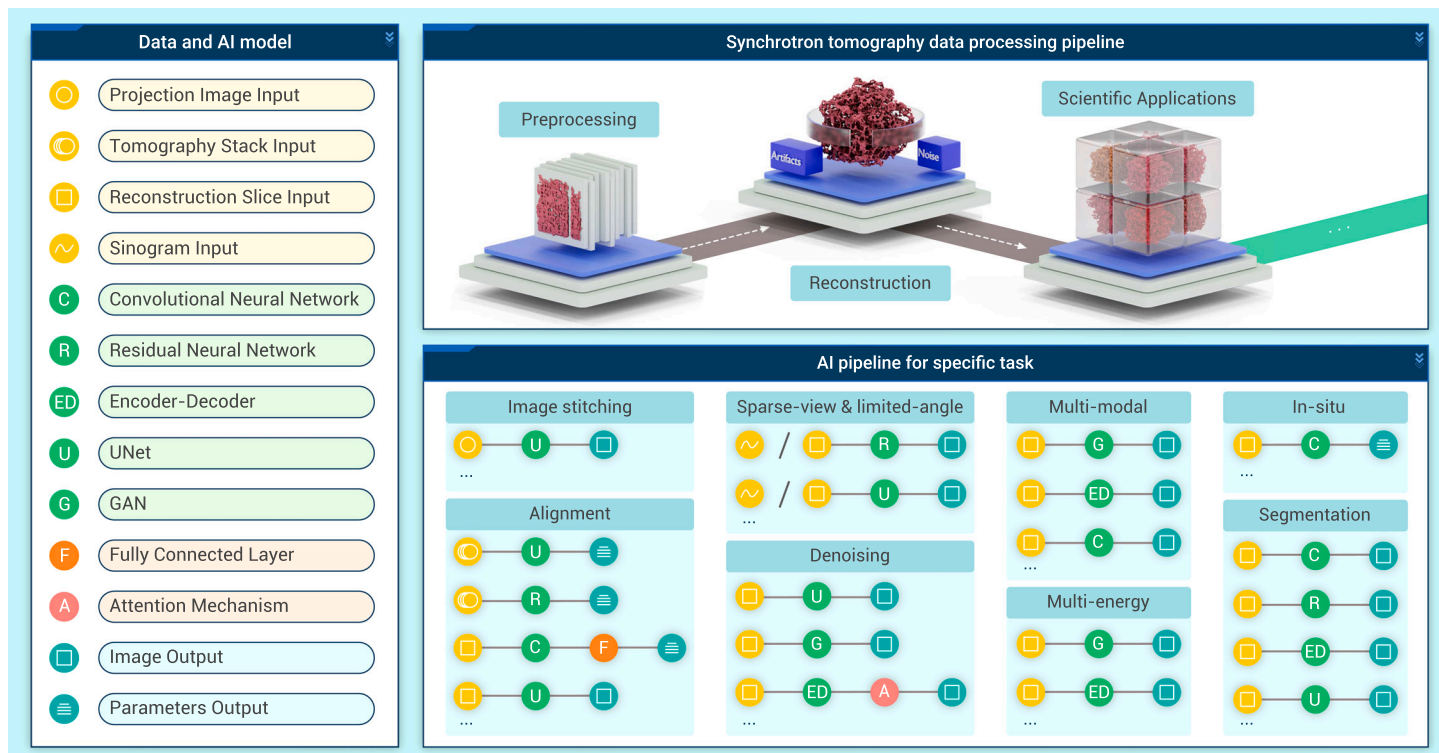


Figure 1. Overview of DL in the STDPP DL is tightly integrated with each part of the STDPP to improve experimental efficiency and accuracy.

two fields, well-established methods will be included in this review. We will first provide an overview of the evolution of DL techniques. Then, we concentrate on DL applications in the preprocessing stage of the STDPP, summarizing DL achievements and drawbacks in this phase. Next, we discuss how DL, combining with techniques used in other scientific fields, can help handle the reconstruction optimization problem. Additionally, we deal with DL applications in the scientific application-oriented data processing tasks after 3D reconstruction, with an emphasis on region of interest (ROI) segmentation tasks, multi-modal tomography, multienergy tomography, and *in situ* tomography. Finally, we discuss prospective challenges, opportunities, and outlooks DL is encountering in the STDPP by proposing three critical components: synchrotron tomography big model, intelligent scheduling center, and single-facility learning and long-term learning.

OVERVIEW OF THE EVOLUTION OF DL TECHNIQUES

DL has been a research hotspot in machine learning since its emergence in the 1980s. Early DL models were often regarded as ANNs with an MLP architecture (Figure 2). However, due to the massive number of learning parameters and the vanishing gradient problems, the stability of ANNs during the training phase could not always be guaranteed, and their performance in image processing tasks often fell short. In 1998, CNNs were first introduced in machine vision applications with the introduction of LeNet.³¹ CNNs typically consist of multiple layers, such as convolutional, pooling, and fully connected (FC) layers, which effectively alleviate the problem of parameter redundancy and training difficulties. In 2012, Krizhevsky et al.³² proposed AlexNet, which has deeper structures and uses rectified linear units (ReLU)³³ as an activation function instead of Sigmoid. The number and quality of extracted features using AlexNet significantly surpassed those using manual approaches, making AlexNet the inception of modern CNNs.

In the following years, numerous meticulously designed CNN architectures emerged³⁴⁻³⁶ with deeper network structures, but training difficulties and the vanishing gradient problem still persist, and the performance is nearly reaching the limit. However, in 2015, He et al.³⁷ proposed ResNet, which dramatically improved the performance of CNNs. ResNet uses residual blocks by summing input and output features to fuse information, circumventing vanishing gradients incurred by overly deep network architectures. Today, residual blocks are the primary DL network modules applied in synchrotron tomography.

Recently, the performance of DL has excelled in natural language processing (NLP) tasks using transformer architectures. In 2020, a research team from Google proposed Vision Transformer (ViT)³⁸ and applied transformer architectures to computer vision applications. Subsequently, the Swin-transformer³⁹ network has pushed the performance limit of CNNs even further. Plus, the Latent Diffusion Model (LDM),⁴⁰ proposed in 2022, achieved a performance breakthrough in image generation tasks using a generative adversarial network (GAN).⁴¹ Through progressive development, the number of parameters in DL networks has dramatically increased, and computer vision is finally entering the big-model era. Even today, newer models are actively proposed, such as ChatGPT and Segment Anything Model (SAM),⁴² and the performance of DL models based on transformers still has not reached its limit,³⁸ with rapid development still ongoing.

From the basic MLP, the prevailing CNN, and highly effective ResNet to the more recent ViT, LDM, and ChatGPT, DL has and will continue to demonstrate its potential at each stage of the STDPP: image processing before reconstruction, 3D reconstruction optimization, and scientific application-oriented data processing.

IMAGE PROCESSING BEFORE RECONSTRUCTION

To begin, we will discuss DL's applications on the preprocessing workloads before 3D reconstruction, which include image stitching within each projection in scanning tomography experiments and image alignment between different projections (Figure 3).

Image stitching

In experiments like scanning transmission X-ray microscopy (STXM), only a small portion of the whole 2D projection is captured for each single exposure due to the limited field of view (FoV), and the scanned regions need to be stitched together to obtain a full-projection image. Direct stitching within the scanning area can cause discontinuities and distortions due to drift and jitter of X-ray beams as well as motion inaccuracy from the sample itself. The conventional correction strategies that are currently used, such as feature-based⁴³ and cross-correlation⁴⁴ methods, limit the accuracy and stability of the stitched images.

In the medical field, similar pixel drifts and image distortions are also present in Optical Coherence Tomography (OCT) retinal vascular images, which suffer from internal jitters within the imaging system and disturbances caused by the patient.

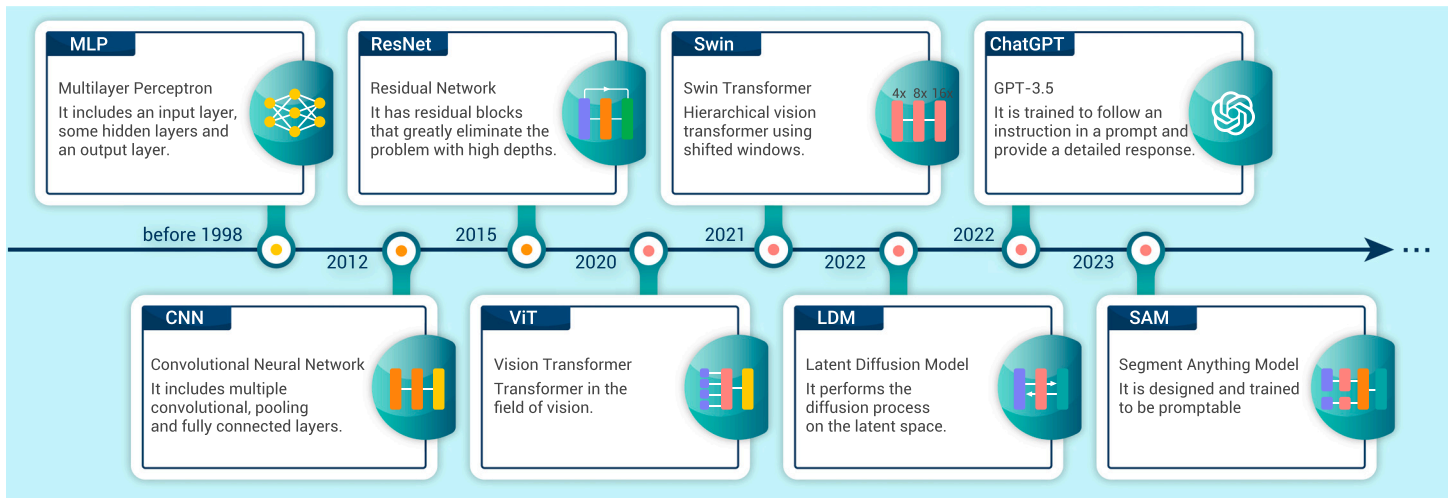


Figure 2. Overview of the development of DL The development of new DL technologies has driven advances in the field of synchrotron tomography.

DL methods have been developed to correct for these distortions in OCT images. Qin et al.⁴⁵ proposed a DL distortion correction network in 2021. They utilized a modified UNet architecture to construct a mapping from distorted images to reference images. The modified model demonstrated competence in complicated tasks such as image distortion recovery. Through supervised learning, distorted images could be corrected to near-ground-truth quality, and similar approaches could be explored for synchrotron tomography. However, unlike OCT images, which have publicly available paired datasets such as OCTA-500,⁴⁶ synchrotron radiation computed tomography (SR-CT) lacks sufficient accessible datasets. Moreover, the sample types examined in SR-CT are much more diverse; hence, the training dataset needs to cover a large variety of sample types. Dataset construction is therefore critical in future distortion correction tasks, particularly when real-world data collection is becoming an arduous and time-consuming task within limited beamtime. Because the mechanisms behind distortion are usually uniform and predictable for specific beamlines, one possible solution for data construction is simulation using well-established distortion mechanisms.

Alignment

Aligning the projection images is another crucial step before reconstruction and aims to correct the misalignments between projection images that can result in artifacts on reconstructed data. Misalignments are generally inevitable in high-resolution tomography experiments due to mechanical instabilities or positioning errors, whether for full-field tomography methods or scanning tomography combined with ptychography, spectroscopy, or diffraction methods. The artifacts are often less obtrusive in microscale synchrotron tomography systems. However, for methodologies performing at nanoscale, the drifting effects would be substantially magnified. When positioning accuracy enters the nanometer regime, the system can no longer be corrected merely by hardware upgrades.

A mechanical system's instability in synchrotron tomography arises from various aspects, such as motion inaccuracy of the sample and drift and vibration of the imaging system during the long acquisition process. These effects are usually coupled in the postprocessing phase, making the alignment task even more demanding. Conventional methods, like center of mass (CoM),⁴⁷ cross-correlation,^{48,49} and iterative techniques,⁵⁰⁻⁵⁴ are often applied for tomography stack optimization, but all have their limitations, mostly on accuracy and efficiency. Correcting coupled or mixed types of motion inaccuracy, drift, and vibration in a faster and more accurate fashion is demanding, and extensive research is being undertaken on DL image alignment in the STDP.

To mitigate instrument instability, in 2016, Yang et al.⁵⁵ proposed a CNN model that calibrates the center of rotation (CoR) in synchrotron tomography. The architecture used was even simpler than that of AlexNet, comprising only two consecutive feature extraction modules. The model distinguishes reconstructions corresponding to correct or incorrect CoRs.⁵⁵ The results matched the performance achieved by human estimation. Most importantly, such a network can be

embedded naturally into a data processing pipeline, enabling automatic correction capabilities on rotation axes.

For the more complicated correction task involving both sample drift and detector position inaccuracies, DL can be applied either independently or jointly by combination with conventional techniques, where DL tackles the image alignment problem with an end-to-end correction scheme. In 2020, Topal et al.⁵⁶ proposed a CNN-based DL network to address the component's thermomechanical instability, where CoR misalignment and detector position inaccuracy happen simultaneously on a ball bearing sample. The network performs segmentation and localization of the same feature-rich regions across different projections. Correction is then made on the CoM using segmented regions. Additionally, in 2022, Fu et al.⁵⁷ proposed a more advanced UNet architecture that segments a common feature area from a series of nano-resolution, full-field transmission X-ray microscopy (nano-CT) stacks acquired from various projection angles. These networks compensate for drift by using the center of a common feature area. Thanks to such advances, estimating sample motion no longer depends on manually added feature points (e.g., using gold particles). Due to the similarity of different instance segmentation problems, advanced instance segmentation networks in computer vision can be naturally migrated into synchrotron tomography. For applications using solely DL methods, in 2021, Fu et al.⁵⁸ proposed a residual network to correct image jitter present in nano-CT. Unlike using original projection images as network input, this model combines the original projections with the reconstructed back-projection images as input. The final output measures the amount of motion compensation of the tomography stack. The reconstructed slices of a battery cathode particle also present better contrast with more detailed internal structures unveiled.

However, due to complex, diverse, and unpredictable sample environments, the generalizability of data processing models in these complicated correction tasks remains a critical concern, mostly owing to the requirement for an immense amount of training data, which can be hardly obtained currently. To improve the model's generalization capability, in 2022, Liu et al.⁵⁹ designed the multiscale dense U-Net. They adopted data augmentation techniques such as pruning, scaling, shifting, and rotating. They also used MIMO-Unet⁶⁰ architecture, which addresses the task of multiscale fusion by using images with distinctive scales as input to provide abundant prior information. The performance and robustness of the entire network were improved by combination with embedded dense blocks.⁶¹ When encountering other sample types, the network can still obtain good alignment results with only a small amount of training data provided. However, the problem of data scarcity encountered in supervised learning still exists, and alternate schemes, such as unsupervised or self-supervised learning, will be more suitable to improve models' generalizability. For instance, the pretext task in contrastive learning⁶²⁻⁶⁶ may be used to replace the laborious manual labeling procedure so that the network can be self-trained successfully.

3D RECONSTRUCTION OPTIMIZATION

We will now discuss DL applications in the 3D reconstruction stage, focusing on three subjects: optimization of 3D reconstruction for sparse-view acquisition,

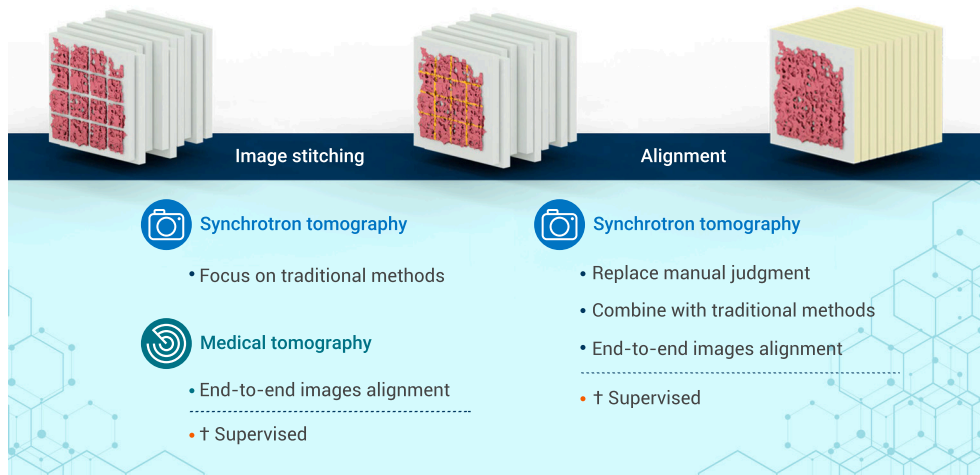


Figure 3. Combination of DL and image processing before reconstruction DL provides excellent pre-processed projection images for 3D reconstruction.

limited-angle acquisition, and denoising on reconstructed volumes or images (Figure 4).

Sparse view

To track the 3D structural and functional transformation of the examined object in *operando* or *in situ* experiments, sparse-view acquisition is often employed to speed up the measurement process or alleviate radiation damage. Due to the insufficient images acquired, streak artifacts (Figure 4, green area in sparse view) are prominent on reconstructed tomography slices. Traditional methods usually include compressed sensing approaches that progressively recover data from sparsely acquired images, which is an iterative and time-consuming process. Hence, DL approaches are being applied to optimize sparse-view synchrotron tomography to better handle the trade-off between accuracy and efficiency. Depending on the data processing goals, these approaches can be divided into optimization in the sinogram domain, optimization in the reconstruction domain, and optimization in both domains.

For optimization in the sinogram domain, a common approach is to use the sinogram⁶⁷⁻⁶⁹ directly as input, which better decouples from the reconstruction process. For example, in 2018, Liang et al.⁶⁷ proposed a deep residual convolutional network to recover the sinogram in the projection domain, significantly reducing streak artifacts in subsequent reconstruction. However, due to the large number of projection angles and pixel counts, taking the entire sinogram as input would dramatically increase the data volume and make the training process more difficult. Therefore, in 2018, Lee et al.⁶⁸ introduced a UNet architecture to obtain a lightweight design. The team divided the sinogram into patches of size 50×50 pixels and separately fed them to the network to alleviate the burden on random access memory (RAM) during training. The trained UNet achieved good results in metrics like structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR). In 2022, Okamoto et al.⁶⁹ divided the sinogram into band patches and devised an even lighter residual network that successfully augmented the sinogram data and suppressed streak artifacts. Using band patches, more vertically detailed information is attained than using 50×50 patches without a significant increase in input volume, which could considerably benefit the sinogram recovery process.

For optimization in the reconstruction domain,⁷⁰⁻⁷² UNet is frequently used. The architectures of these UNets are similar. One classic network, FBPCovNet, was introduced by Jin et al.⁷⁰ in 2017. FBPCovNet takes one of the slices acquired from filtered back-projection (FBP) as input, and through four downsampling and upsampling layers, optimized slices are obtained. The cross-entropy of the optimized slices from the network output and the ground truth is used as the loss function, and then supervised training is performed. In addition to UNet, in 2022, Okamoto et al.⁷¹ performed streak artifact removal for reconstructed slices using a specialized CNN that incorporates residual blocks and inception blocks³⁶ for better feature fusion capability. This model can recover reconstructed images obtained from sparse-view tomography and can also be used for reconstructed image optimization tasks under cone-shaped incidence beams. Furthermore, rather than directly obtaining optimized images, in 2018, Xie et al.⁷² contrived a GoogLeNet-based artifact-learning network. The optimized image is obtained by subtracting the artifacts

of the learned image from the original image, similar to Noise2Noise.⁷³ Due to the similarities of band artifacts on reconstructed slices, unsupervised methods can be applied to help learn these similarities among slices, remove image artifacts, and improve the quality of subsequent reconstructed images.

For optimization in both domains, sinogram and reconstructed slices can be jointly studied for projection images in sparse-view tomography. In 2023, Gao et al.⁷⁴ proposed an attention-based dual branch (CT branch and sinogram branch) network called ADB-Net. For the sinogram branch, atrous spatial pyramid pooling (ASPP) combined with convolutional layers is applied on sinusoidal images for higher-level feature extraction globally. For the CT branch, the features are fused with the extracted features obtained from the sinogram branch using an attention-based approach through downsampling in a UNet architecture. The strategy strengthens the information critical to sparse-view optimization while weakening the less relevant information. Compared with the optimization technique using a single domain,⁷⁰ ADB-Net showed superior results in terms of root-mean-square error (RMSE), SSIM, and PSNR due to the additional domain. However, due to limitations of the training dataset, ADB-Net still needs to be fine-tuned to obtain optimal performance on new input data. Also, the number of network parameters is enormous because the sinogram utilizes fully connected layers for feature extraction.

In summary, DL is being used for optimization in the sinogram domain, in the reconstruction domain, and in both domains with improved accuracy. The similarity between optimizations of sparse-view tomographic reconstruction and compressed sensing in computer science implies that thoroughly tested models can be directly adopted into synchrotron tomography without major modification.

Limited angle

During the data acquisition process, certain factors can prevent the acquisition of a full 180° tomography stack at synchrotron beamlines. Such factors include constrained angles arising from sample stage and mechanical facilities or from intrinsic sample characteristics, resulting in limited-angle acquisition. Such acquisition leads to information loss in the wedge-shaped area during tomography reconstruction in the Fourier domain, and the reconstructed slices will also have wedge-shaped artifacts. Effective reconstruction algorithms are therefore needed to recover the reconstructed slices containing these artifacts.

Traditional methodologies, such as the algebraic reconstruction technique (ART)⁷⁵ and simultaneous iterative reconstruction technology (SIRT),⁷⁶ are widely used in limited-angle synchrotron tomography and generally outperform the FBP algorithm. However, with a slow reconstruction process, the drawback of the ART and SIRT algorithms is often their computational efficiency. Therefore, DL approaches are being used to improve reconstruction efficiency through supervised, unsupervised, or self-supervised learning strategies. Exhaustive studies have demonstrated the effectiveness of limited-angle synchrotron tomographic reconstruction using UNet and GAN architectures.⁷⁷⁻⁸¹

For supervised DL network models, the optimization problem involves domains of reconstructed slices and sinograms. The cross-entropy between the optimized output slices and the ground truth is usually considered as the loss function during the training phase. Reconstruction algorithms, such as FBP^{70,77} or Simultaneous Algebraic Reconstruction Technique (SART),^{78,79} can be used with these models, and it is critical to determine which algorithm yields better optimization results. SARTConvNet, a performant network proposed by Wang et al.⁷⁸ uses the SART algorithm for sinogram reconstruction as well as a classic UNet architecture with BatchNorm added for regularization. This network outperformed other methods^{70,80} in terms of PSNR and SSIM. DL can



Figure 4. Combination of DL and 3D reconstruction optimization DL is used to improve the quality of reconstructed slices, especially in scenarios where data acquisition is impacted by factors such as sparse view, limited angle, and high noise.

also learn to reconstruct optimized slices from a sinogram, where the loss function utilizes the cross-entropy information between the output sinogram and the ground truth during training. GANrec, proposed by Yang et al.⁸¹ in 2020, is based on GAN architecture and uses a generator network to learn the reconstruction and optimization process. The algorithm was verified on a simulation model extracted from the 3D structure of a shale sample, showing much improved reconstruction accuracy using SSIM and PSNR metrics. In the field of ET, due to limitations regarding the rotation range of the sample stage, supervised DL optimization approaches^{82,83} on reconstructed slices or sinograms have also been studied, with UNet architecture widely adopted as the generator model of GAN. Wang et al.⁸² effectively recovered data with only a range of -50° to 50° using UNet++ as the generator of GAN, while Xin et al.⁸³ achieved good recovery results on missing ET data at 45° by combining reconstructed slices and sinograms using UNet as the GAN generator.

Because paired datasets are not always readily available, and the amount of usable data is often limited, the generalizability of DL networks is becoming a concern. Unsupervised learning can solve the challenge of collecting labeled data and expanding data volume, which could improve generalizability using a massive amount of expanded training data. In 2021, Barutcu et al.⁸⁴ extended the functionalities of GANrec and developed the TomoDIP_TV network by combining the unsupervised deep image prior (DIP).⁸⁵ TomoDIP_TV transforms the reconstruction model at limited angles into a constraint optimization on auxiliary variables during the reconstruction process. The network uses FC layers and 3D convolutions for feature extraction, and the output 3D reconstructed object is obtained following the last convolutional layer. Back-projection is then performed to update the auxiliary parameters until the specified number of iterations is reached. Using DIP to improve the reconstruction, the network implicitly optimizes the weights that recover the missing parts of projections, with assistance from a physical forward model. Compared with GANrec, the network also demonstrates superior robustness to noise. However, due to the self-training strategy used, the model requires another complete training process whenever new data are presented, which is much less efficient than supervised learning methods. Additionally, the implementation of 3D convolutions imposes a significant burden on computational resources.

Denoising

Denoising on reconstructed images is often considered the most common task in 3D reconstruction optimization. In most *in situ* and dynamic tomography experiments, the exposure time to acquire each projection image is reduced significantly if sparse-view acquisition is not applied, resulting in lower-SNR images that may further decrease reconstruction quality. Therefore, denoising techniques should be applied on reconstructed images to enhance the quality of recovered structures. Similarly, in the medical field, concerns about radiation dose also exist due to the involvement of human subjects. The progress⁸⁶⁻⁸⁹ made so far by researchers in the medical field is expected to benefit synchrotron tomography.

The network models for denoising reconstructed synchrotron tomography images have evolved from supervised⁹⁰⁻⁹⁴ to self-supervised⁹⁵ net-

works. One of the representative works on supervised denoising is TomoGAN, which was proposed by Liu et al.⁹⁰ in 2020. They adjusted the GAN architecture to focus on finely detailed features in high-frequency content of synchrotron tomography. The generator follows the UNet structure for biomedical image segmentation.⁹¹ The network was trained using simulated data and verified with both simulated and real-world images, demonstrating significant improvements in SSIM and computation time compared with iterative SIRT algorithms. In addition, in 2023, Yang et al.⁹² applied transformer's multihead self-attention mechanism to retrieve sinogram features from different viewing angles owing to its capability of extracting interrelation between sequence data. They trained a residual network to remove artifacts using the denoised sinogram and noise-corrupted reconstruction slices. The method achieved better results than iterative reconstruction methods and CNN models,^{88,93} even at extremely low (5%) dose levels.

Collecting noisy-clean tomographic image pairs can be laborious or even impractical, making self-supervised denoising networks a research hotspot. In 2021, Hendriksen et al.⁹⁵ proposed Noise2Inverse, which uses a CNN to remove noise from reconstructed images without a paired dataset. The method divides each sinogram into subsinograms with equally spaced angles to complete noise learning and denoising. It considers the similarities between the noise distributions of the reconstructed image from one subsinogram and the slices reconstructed from the remaining subsinograms, all using the FBP reconstruction algorithm. The network was compared with SIRT and total variation minimization (TV-MIN) on the TomoBank⁹⁶ experimental dataset, demonstrating superior performance in terms of SSIM and PSNR under various noisy configurations. This study demonstrates the capability of self-supervised learning in denoising synchrotron tomography images.

In the medical area, the challenge of constructing paired noisy-clean images still exists. Hence, self-supervised learning methods have also been well studied. In 2020, Qiu et al.⁸⁶ proposed N2NSR-OCT, which utilizes Noise2Noise to learn the similarities between noise distributions of paired training data, including original OCT images and cropped images. In 2022, Jing et al.⁸⁷ studied a supervised learning framework on low-dose CT (LDCT) denoising using self-attention (SA) modules. The framework does not require any normal-dose CT (NDCT) images as labels and includes SA modules in the architecture of residual encoder-decoder convolutional neural network (RED-CNN),⁸⁸ an encoder-decoder network. In 2023, Selim et al.⁸⁹ used a diffusion model for the unsupervised noise reduction task of LDCT images to achieve highly accurate noise-free image generation with significantly improved runtime.

In summary, several significant studies have been done on denoising synchrotron tomographic reconstruction images. The similarities between synchrotron tomography, medical imaging, and computer vision have allowed the migration of advanced denoising models and algorithms from computer vision into the denoising task on synchrotron tomography reconstruction. Widely adopted models, such as CNN or GAN, have been used, and other models, like Swin-Transformer, may be considered as a more performant baseline for feature extraction. The diffusion model can also be used to surpass the performance

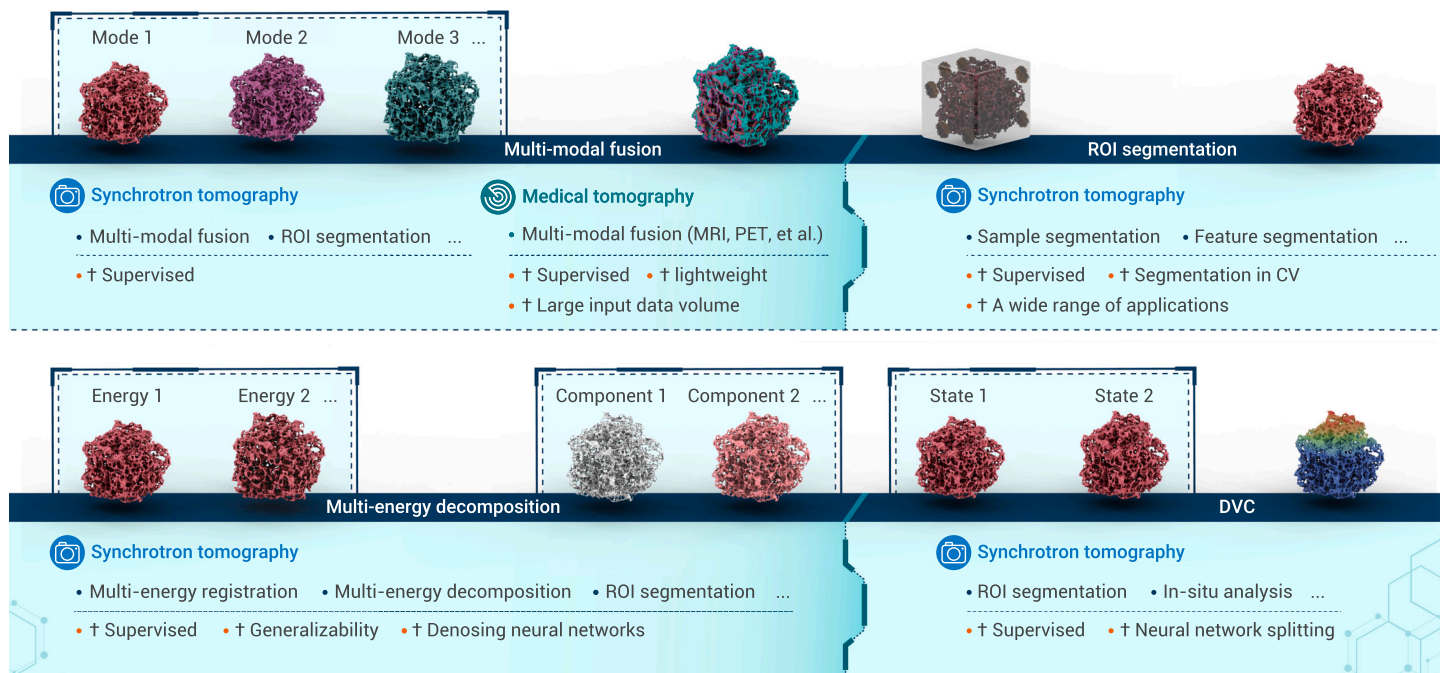


Figure 5. Combination of DL and scientific application-oriented data processing on reconstructed data The versatility of scientific goals leads to a blossoming of DL applications in synchrotron tomography.

of GAN. In the future, these models could bring about impressive denoising results in the field of synchrotron tomography.

SCIENTIFIC APPLICATION-ORIENTED DATA PROCESSING ON RECONSTRUCTED DATA

Following 3D reconstruction, the scientific application-oriented data processing tasks are closely linked to real-world scientific goals. In the following, we will examine the applications of DL in four areas of scientific application-oriented data processing: ROI segmentation, multimodal analysis, multienergy analysis, and *in situ* tomography (Figure 5).

ROI segmentation

ROI segmentation is often considered one of the most common scientific application-oriented data processing tasks in the STDPP. The ROI segmentation task is frequently encountered, whether in single-modal or more complex multi-modal and *in situ* synchrotron tomography experiments. Due to its similarities to computer vision applications like instance segmentation, DL-based ROI segmentation in synchrotron tomography has been thoroughly studied. Depending on the objects to be segmented, ROI segmentation can be divided into reconstructed intra-voxel segmentation and feature segmentation inside samples.

In some experiments of synchrotron tomography, the true structural distributions are difficult to observe and analyze due to degraded imaging contrast and disturbance from ROIs. Therefore, the segmentation task aims to accurately identify the voxels belonging to highly complex sample structures or partial areas of interest from reconstructed 3D voxels, which would aid in the subsequent analysis process. Numerous efforts have been made, including the use of CNN-based models,⁹⁷ UNet,^{98,99} and Mask R-CNN.¹⁰⁰ Here, we focus on the UNet-based approach and the highly accurate Mask R-CNN approach. In 2022, Davydenko et al.⁹⁸ introduced a typical UNet architecture for highly accurate segmentation of complex Mg-based alloys among reconstructed voxels. In addition to the segmentation advantage of the UNet architecture itself, segmentation results were further improved using a data augmentation method suitable for the tested samples. The augmentation process utilizes randomized sampling and cross-correlation filling methods to expand the current small dataset constrained by a limited number of segmenting labels. For segmentations of composite materials, such a method is efficient and effective because no stringent requirements are raised regarding the continuity of the material itself. Additionally, in 2021, Torbati-Sarraf et al.⁹⁹ performed a comprehensive comparison among UNet, UNet++, PSPNet, and DeepLab v3+ network models. Taking the complex structures of 7075 aluminum alloy (AA7075) as an example, the effectiveness of auto-

matic segmentation on transmission X-ray microscopy (TXM) tomographic data using various models was validated. The basic UNet yields the lowest training time with mediocre performance, while the performance greatly improved for the more complex UNet++¹⁰¹ model, which incorporates skip connections. PSPNet¹⁰² and DeepLab v3+¹⁰³ were pretrained using ImageNet¹⁰⁴ and achieved decent results, but the performance was still limited. It has been proven that advanced models used for instance segmentation in computer vision can improve the segmentation performance in the field of synchrotron tomography.

For Mask R-CNN,¹⁰⁵ excellent performance in instance segmentation has been demonstrated in the field of computer vision. For synchrotron tomography, Lin et al.¹⁰⁰ implemented the Mask R-CNN model for the segmentation of vanadium pentoxide (V_2O_5) nanoparticles in 2022. The team encountered the problem of insufficient training data and developed a randomized nanowire generator using Geodict for data generation. They successfully obtained a large amount of simulated training data. Due to the certainties of V_2O_5 structures, the model exhibits superior segmentation capabilities on real-world multimodal datasets obtained from X-ray ptychography, STXM, and scanning electron microscopy (SEM) experiments.

In the field of ET, segmentation of the target particle within the imaging area is a significant challenge due to complex cellular structures and the impact of incorporated ice. For DL architectures, common models, like CNN¹⁰⁶ and encoder-decoder,¹⁰⁷⁻¹¹³ are often used. To segment a single particle, some researchers^{106,108} combined prior knowledge with these DL models to get more easily trained networks. In addition, research on image segmentation using a diffusion model (DM) is being carried out in the medical field. In 2023, Zhang et al.¹¹⁴ proposed a generalized hybrid denoising DM (GH-DDM) for cryo-ET images to improve the quality of medical image generation in the absence of high-quality labeled data. Wu et al.^{115,116} proposed MedSegDiff and MedSegDiff-V2, both with better final testing performance compared with the state of the art (SOTA).

For the task of feature segmentation inside samples, the focus is on segmenting evolving changes in *in situ* synchrotron tomography. Such changes may include features not belonging to intra-voxel information, such as gaps. UNet-based^{117,118} models are often applied here. In 2021, Kopp et al.¹¹⁷ implemented a 5-layer downsampling network architecture for segmenting various types of microscale damage in heterogeneous bulk materials. Additionally, in 2022, Fu et al.¹¹⁸ successfully segmented microcracks in an *in situ* lithium-ion battery using a standard 4-layer downsampling UNet. Compared with the 65,000 images required by Kopp et al.,¹¹⁷ the team only had 48 images for

training and 16 images for validation. Therefore, they expanded the dataset using data augmentation techniques such as rotation, scaling, and cropping. In synchrotron tomography, it is often difficult to attain an immense dataset matching research objectives due to sample complexities. Thus, data augmentation techniques¹¹⁹ that expand the size of small datasets are commonly considered effective.

Multimodal analysis

As the experimental mode for synchrotron tomography is shifting toward multimodal characterization, data fusion between multiple modalities is becoming one of the most challenging steps in the entire STDP. For example, the fusion of 3D tomography datasets is becoming more difficult due to data throughput explosion at next-generation synchrotron light sources. In synchrotron tomography, conventional and machine learning methods^{120,121} are still considered the most popular, while DL approaches from other data-intensive scientific fields are also worth investigating. GAN-based architectures have been found to be effective^{122,123} in performing registration and correlation from different modalities because the end-to-end generative nature of multimodal tomographic data fusion workload is well suited. In 2022, Abirami et al.¹²² used GAN to successfully register single-layer slices obtained from magnetic resonance imaging (MRI) and positron emission tomography (PET) scans, achieving excellent contrast on soft tissues using MRI and better spatial characteristics on bones using PET. The advantages of the two modalities are effectively complemented through modality fusion.

The spatial resolution and FoV for each modality may vary when the tomography datasets are not collected simultaneously, leading to a cross-scale data registration and correlation demand on top of the data fusion. In 2022, Liu and Mukerji¹²³ performed data fusion on carbonate samples obtained by SEM and micro-CT to better understand liquid flows at the porosity level. The structural information covering a wider FoV is obtained from micro-CT, while the finer pore-scale features within a narrower FoV are acquired through SEM. The network used in the study consists of a StyleGAN2¹²⁴ network and a CycleGAN¹²⁵ network. StyleGAN2 generated SEM images close to the ground truth from randomized noise and the style latent vector, while CycleGAN consists of two pairs of generators and discriminators. From the low-resolution micro-CT images, the first pair creates reconstructed images close to the SEM images obtained by StyleGAN; from the reconstructed high-resolution SEM images, the second pair produces reconstructed images close to the low-resolution micro-CT images, which ensures cycle consistency. Eventually, the images from the two scales are fused together using the images from CycleGAN's SEM generators and micro-CT generators.

It is important to note that collecting paired datasets for multimodal fusion tasks is always considered a challenging task. However, obtaining labels for segmentation is often easier. Therefore, in 2019, Blendowski et al.¹²⁶ transformed the problem of directly fusing multimodal images into the problem of linear interpolation of the shared features across modalities. They used an encoder-decoder network consisting of 3D convolutions to extract information within the shared features of the reconstructed 3D images from various modalities. Later, linear interpolation was applied to complete correlation and registration on a broader scale. Although limited by segmentation quality, the proposed method successfully addressed the problem of paired dataset collection in the complicated multimodal data fusion task. Furthermore, due to the large number of computing resources required by 3D convolutions, which hinders network training and subsequent deployment, Hering et al.¹²⁷ proposed 2.5D convolutions as a substitute for 3D convolutions. The basic idea is to segment the reconstructed voxels into slices oriented toward three directions and compute 3D compensations after proper registration in these directions. This method dramatically reduced the burden on computing resources while guaranteeing registration accuracy.

Multienery analysis

Multienery synchrotron tomography provides extra differentiation capability on material compositions by exploiting the energy dependence of X-ray attenuation for different materials,¹²⁸ following registration and decomposition. In such workloads, the magnification ratios used in tomographic and multimodal tomographic registration tasks are similar because they both involve registration of different magnification ratios, angles, and characterizations. Therefore, the

network design should also be similar. Registrations can be performed using UNet or encoder-decoder architectures. For the material decomposition, conventional least-squares methods under nonnegative constraints are often applied. Such methods can incur significant noise. Therefore, DL has been adopted to study the decomposition of different materials with high fidelity.

In 2019, Clark et al.¹²⁸ used a simple 2-stage downsampling UNet to separate materials from multienery CT images. Due to the lack of a real-world dataset, simulated datasets containing iodine, barium, photoelectric effect, and Compton scattering were used to construct the paired training data for supervised training. Using real datasets, the trained network “smoothly” separated iodine and barium with lower noise levels, which would be more difficult using traditional methods. In addition to the basic UNet, deeper networks were also implemented^{129,130} to further improve decomposition accuracy. For example, in 2020, Gong et al.¹²⁹ applied an encoder-decoder network containing inception blocks.³⁶ They retained more preserved structural details while further suppressing noise levels. However, neither network could competently deal with the generalization problem caused by data scarcity. We believe that this problem can be better addressed when the unsupervised learning approaches mentioned under **alignment and denoising** are considered.

Last, because the challenge induced by noise cannot be effectively addressed solely through conventional methods, denoising the decomposed images using DL could provide an effective solution. Since 2020, Fang et al.¹³¹ have performed denoising directly on decomposed multienery data from different materials. They successfully separated distinct materials using the Noise2Noise approach, achieving satisfactory results in terms of detail recovery as well as noise suppression.

In situ tomography

The capability of fast tomography data acquisition in new-generation synchrotron beamlines allows *in situ* characterization of 3D structural and functional evolution processes of samples under various effects caused by force, heat, electricity, magnetic environmental changes, etc. Especially in materials science,¹³²⁻¹³⁴ it is critical to quantitatively analyze structural deformation and strain information from reconstructed datasets or segmented features^{117,118} to identify key functional mechanisms, which turns out to be technically challenging and computationally expensive using conventional algorithms. For example, traditional solutions to 3D deformation characterization usually include the digital volume correlation (DVC)¹³⁵ technique, which calculates internal 3D displacement and strain fields during the deformation process. The technique has been widely adopted in the past 20 years, but due to the numerical iterative process, the computational cost of the entire DVC procedure can still be significant, even with graphics processing unit (GPU) acceleration.¹³⁶ Using the DL solution to address such time-consuming task has just begun. For example, in 2022, Duan and Huang¹³⁷ proposed DVC-Net, which specifically tackles the time-consuming problem caused by numerical iteration. Due to the complexity of DVC, DVC-Net was divided into three subnetworks, with each trained separately. Each subnetwork is a simplified CNN architecture that, respectively, extracts the whole voxel deformation, subvoxel deformation, and “smoothly” acquired displacement fields. The idea of dividing convoluted tasks into subnetworks is noteworthy. Subnetworks present numerous advantages. For instance, training separately using subnetworks could save computing resources, and the problem's objectives would become more specific with simpler model structures, which benefits the training process. Additionally, collected datasets from other fields highly relate to specific scientific objectives, which would also aid in training. As for results, through the division and combination of three subnetworks, DVC-Net successfully achieved accurate and robust DVC calculations, expediting runtime by two to three orders of magnitude.

CHALLENGES, OPPORTUNITIES, AND OUTLOOK

DL is already being widely studied and adopted in the field of synchrotron tomography. It is being used for various tasks, such as distortion correction in image stitching before 3D reconstruction; preprocessing tasks, such as instability mitigation across domains; artifact removal in sparse-view and limited-angle imaging; and 3D reconstruction optimization, such as denoising induced by underexposure; it is also used for complicated and various scientific application-oriented data processing tasks, including ROI segmentation and multimodal, multienery, and *in situ* tomography. DL has already demonstrated its superiority

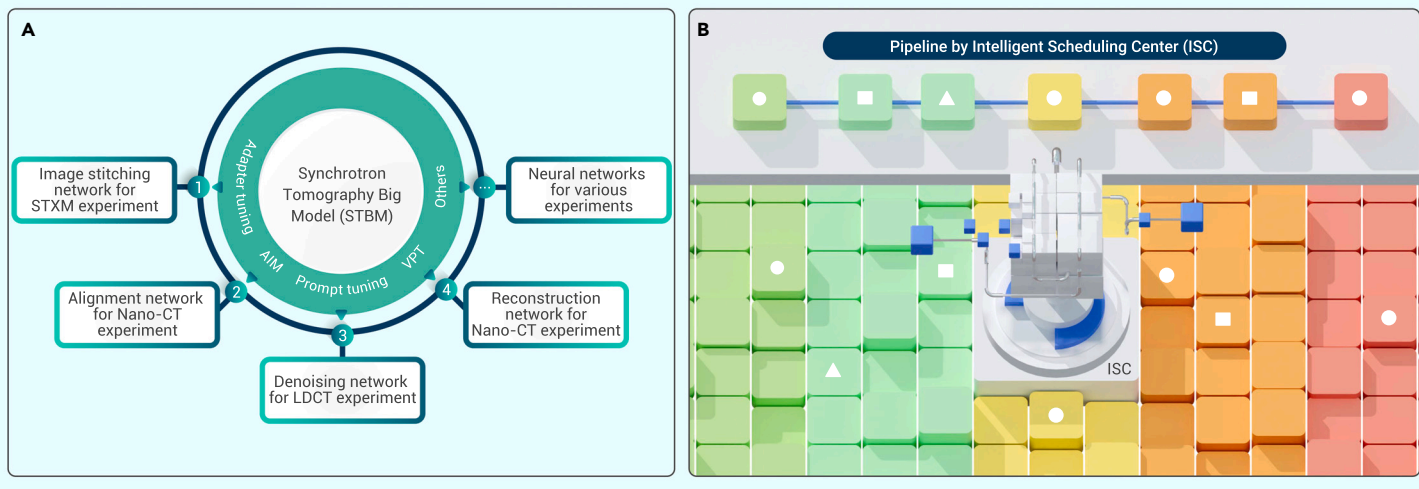


Figure 6. Outlook for DL in synchrotron tomography (A) Different types of networks in downstream tasks are derived through the STBM, using advanced tuning methods. (B) The ISC makes selections from numerous functional DL models to formulate the STDPP.

in data processing accuracy and efficiency and is expected to gradually replace conventional paradigms in the future.

It should be noted that, due to sample variability and lack of paired data, the feature extraction capability of DL models directly trained from the input end would be limited, resulting in poor accuracy and generalizability in the downstream tasks of the STDPP. Consequently, backbone networks that specifically address the generalizability problem in synchrotron tomography are essential, which is referred to as the synchrotron tomography big model (STBM) (Figure 6A). Using a pretrained STBM in concert with customized downstream networks, better feature extraction results could be achieved for various tasks in the STDPP, making the training process more accurate with less computing resources consumed. For synchrotron beamlines often limited by computing power, STBM is expected to facilitate the development and applications of DL in the STDPP. However, it is still critical to collect an immense amount of data and devise proper network architectures for the STBM. Because synchrotron tomography data are usually categorized into three domains (projection, sinogram, and reconstruction), the corresponding STBMs can be divided into STBM-P, STBM-S, and STBM-R, which all require the acquisition of a large amount of training data in each domain to achieve decent generalizability. Hence, the construction of the training dataset would be more straightforward. In terms of model design, because the ability to learn from a tremendous amount of data is a requisite for the STBM, model prototypes such as ViT, Swin-Transformer, and LDM are considered competent candidates. Moreover, for synchrotron tomography, the main objective of STBM design is feature extraction. Therefore, the training goal could be more flexible with concentration on low-level tasks such as classification, which could eventually ease the training process.

Similar to the big models in computer vision, which contain billions of network parameters, for downstream tasks, directly fine-tuning the entire network using a small dataset could instead downgrade the generalizability and accuracy, let alone the tremendous computational cost induced by a large number of model parameters, which contradicts the initial motivation when designing STBMs. Therefore, the parameters of the STBM should be kept intact when possible. For distinctive downstream tasks, with the parameters of the backbone networks frozen, parameter-efficient fine-tuning (PEFT)¹³⁸ could be applied for training, such as adapter¹³⁹ and adapting image models (AIM).¹⁴⁰ These plug-and-play models require very few training parameters, making the STBM backbone more suitable for customized downstream tasks in the STDPP without redundant modification. Alternatively, prompt tuning (PT)¹⁴¹ and visual PT (VPT)¹⁴² can be used to fine-tune the STBM by incorporating only a small number of training parameters into the input space after freezing the backbone network parameters. It is also worth noting that soft X-ray tomography based on highly coherent rings and ptychography techniques is also an exciting area that involves more dimensions and richer scientific scenarios, leading to more demand in its data processing pipeline. These tuning techniques enable the derivation of downstream STBM-like models for the soft X-ray tomography¹⁴³ pipeline, which

greatly helps reveal multidimensional properties of materials pertaining to structural, chemical, elemental and electronic information.

On the other hand, as the types of models and data scales continue to increase in the STDPP, proper allocation of computing power, as well as the selection of target functional models, is becoming critical for deployment workloads at future large scientific research facilities, such as synchrotron light sources. Limited by the availability, capacity, and capability of computing resources, it is more imperative than ever to create an intelligent scheduling center (ISC) that intelligently allocates computing resources and automatically selects the most suitable models for certain workloads (Figure 6B). Using a pretrained ISC, models residing in the DL library could be autonomously selected and recommended through each stage of the STDPP, which enables the assignment of appropriate pipelines according to distinctive data types and structures for efficient processing and real-time feedback. Also, because the ISC inherently takes into account intermodule correlation and optimal workflows, a graph neural network (GNN) may be used to strengthen such correlation. For optimal workflows, deep reinforcement learning (DRL) can also be applied to formulate the necessary optimal decision strategy. In the training process of DRL, the reward design needs to balance the trade-off between computing power consumption and final tomographic results to reach a “sweet spot.”

Furthermore, a future ISC could be a multimodal model with inputs consisting of images (synchrotron tomography stack) or text (task requirements/metrics) and with outputs consisting of selected pipelines or models. The architecture of contrastive language-image pre-training (CLIP)¹⁴⁴ could also be used to jointly train a universal ISC big model with data from natural language and machine vision. Later, by making use of the Reinforcement Learning from Human Feedback (RLHF)¹⁴⁵ technique that is adopted by ChatGPT, the pretrained ISC model should become more adapted to the scientifically more prudent synchrotron tomography area. Like the STBM, finalized optimal designs of the selected pipelines would differ from each other due to differences in task objectives across various beamlines. Therefore, the ISC also needs to be adapted accordingly based on downstream tasks. It would be feasible to adjust ISC models using the “fourth paradigm” of NLP, such as prompt learning, to meet the pipeline design requirements from different beamlines. The goal is to offer a better user experience to synchrotron tomography users and staff and ultimately improve user services in large scientific research facilities.

Besides experimental data, metadata are also worth studying. Metadata can help build single-facility learning schemes for beamline scientists. For example, collected metadata can be used for training using inherent drift data generated at certain synchrotron tomography beamlines. Additionally, due to unpredictable factors, such as mechanical abrasion, the distribution of the objectives to be optimized will gradually change, where a long-term single-facility learning strategy is more desired. By making use of the collected metadata during the elongated experimental process, DL and DRL methods may be applied to learn the trending changes of objective distribution to fit or predict metric variations at synchrotron

tomography beamlines. Hence, the STDP at certain beamlines can be better optimized. Also, enabled by the model's prediction capability, the system can provide early warnings for beamline instrumentation, such as excessive drift caused by mechanical abrasion.

All in all, whether it is for certain tasks in the STDP, for the STBM, for an ISC, or for scientific application-oriented data processing on reconstructed tomographic data using single-facility and long-term learning, impacted by today's data-driven "fourth paradigm," the full-stack DL-empowered data processing pipeline for synchrotron tomography experiments (STDP) will be facing unprecedented challenges and opportunities but with an inspiring outlook ahead. The STDP has all of the potential of being built, deployed, and refined in tomography beamlines, by tomography methodologies, and for tomography experiments at new-generation light sources in the very near future.

REFERENCES

- Walton, T., Gui, M., Velkova, S., et al. (2023). Axonemal structures reveal mechanoregulatory and disease mechanisms. *Nature* **618**, 625–633.
- Wu, J., Lu, Z., Jiang, D., et al. (2021). Iterative tomography with digital adaptive optics permits hour-long intravital observation of 3D subcellular dynamics at millisecond scale. *Cell* **184**, 3318–3332.e17.
- Talanda, M., Fernandez, V., Panciroli, E., et al. (2022). Synchrotron tomography of a stem lizard elucidates early squamate anatomy. *Nature* **611**, 99–104.
- Mangala Prasad, V., Leaman, D.P., Lovendahl, K.N., et al. (2022). Cryo-ET of Env on intact HIV virions reveals structural variation and positioning on the Gag lattice. *Cell* **185**, 641–653.e17.
- Su, Q., Chen, M., Shi, Y., et al. (2022). Cryo-EM structure of the human IgM B cell receptor. *Science* **377**, 875–880.
- Wozny, M.R., Di Luca, A., Morado, D.R., et al. (2023). In situ architecture of the ER-mitochondria encounter structure. *Nature* **618**, 188–192.
- Domínguez-Martín, M.A., Sauer, P.V., Kirst, H., et al. (2022). Structures of a phycobilisome in light-harvesting and photoprotected states. *Nature* **609**, 835–845.
- Watanabe, R., Buschauer, R., Böhring, J., et al. (2020). The in situ structure of parkinson's disease-linked LRRK2. *Cell* **182**, 1508–1518.e16.
- Zhang, K., Liu, X., Shen, J., et al. (2020). Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* **181**, 1423–1433.e11.
- Ng, K.W., Boumelha, J., Enfield, K.S.S., et al. (2023). Antibodies against endogenous retroviruses promote lung cancer immunotherapy. *Nature* **616**, 563–573.
- Lee, J.H., Liu, Q., and Dadgar-Kiani, E. (2022). Solving brain circuit function and dysfunction with computational modeling and optogenetic fMRI. *Science* **378**, 493–499.
- Han, M., Bushong, E.A., Segawa, M., et al. (2023). Spatial mapping of mitochondrial networks and bioenergetics in lung cancer. *Nature* **615**, 712–719.
- Scott, M.C., Chen, C.-C., Mecklenburg, M., et al. (2012). Electron tomography at 2.4-angstrom resolution. *Nature* **483**, 444–447.
- Wang, L., Liu, T., Wu, T., et al. (2022). Strain-retardant coherent perovskite phase stabilized Ni-rich cathode. *Nature* **611**, 61–67.
- Huang, Y., Perlmutter, D., Fei-Huei Su, A., et al. (2023). Detecting lithium plating dynamics in a solid-state battery with operando X-ray computed tomography using machine learning. *NPJ Comput. Mater.* **9**, 93.
- Li, Y., Yuan, G., Li, L., et al. (2023). Ductile 2-GPa steels with hierarchical substructure. *Science* **379**, 168–173.
- Possenti, E., Conti, C., Gatta, G.D., et al. (2022). Synchrotron X-ray diffraction computed tomography to non-destructively study inorganic treatments for stone conservation. *iScience* **25**, 105112.
- Michelson, A., Minevich, B., Emamy, H., et al. (2022). Three-dimensional visualization of nanoparticle lattices and multimaterial frameworks. *Science* **376**, 203–207.
- Jiao, Y., Xu, G., Cui, X.H., et al. (2018). The HEPS project. *J. Synchrotron Radiat.* **25**, 1611–1618.
- Nazaretski, E., Yan, H., Lauer, K., et al. (2017). Design and performance of an X-ray scanning microscope at the Hard X-ray Nanoprobe beamline of NSLS-II. *J. Synchrotron Radiat.* **24**, 1113–1119.
- Li, X., Zhang, Y., Liu, Y., et al. (2023). A high-throughput big-data orchestration and processing system for the High Energy Photon Source. *J. Synchrotron Radiat.* **30**, 1086–1091.
- Pfeiffer, F. (2018). X-ray ptychography. *Nat. Photonics* **12**, 9–17.
- Kuan, A.T., Phelps, J.S., Thomas, L.A., et al. (2020). Dense neuronal reconstruction through X-ray holographic nano-tomography. *Nat. Neurosci.* **23**, 1637–1643.
- da Silva, J.C., Pacureanu, A., Yang, Y., et al. (2017). High-energy cryo X-ray nanoimaging at the ID16A beamline of ESRF. *SPIE: X-Ray Nanoimaging: Instruments and Methods III*, 10389.
- Walsh, C.L., Tafforeau, P., Wagner, W.L., et al. (2021). Imaging intact human organs with local resolution of cellular structures using hierarchical phase-contrast tomography. *Nat. Methods* **18**, 1532–1541.
- Nazaretski, E., Yan, H., Lauer, K., et al. (2017). Design and performance of an X-ray scanning microscope at the Hard X-ray Nanoprobe beamline of NSLS-II. *J. Synchrotron Radiat.* **24**, 1113–1119.
- Schroer, C.G., Baumbach, C., Doehrmann, R., et al. (2015). Hard X-ray nanoprobe of beamline P06 at PETRA III. *NIM-A* **616**, 93–97.
- Stampanoni, M., Groso, A., Isenegger, A., et al. (2006). Trends in synchrotron-based tomographic imaging: the SLS experience. *SPIE: Developments in X-Ray Tomography V*, 6318.
- Dong, Y., Li, C., Zhang, Y., et al. (2022). Exascale image processing for next-generation beamlines in advanced light sources. *Nat. Rev. Phys.* **4**, 427–428.
- Kuhn, T.S. (1962). *The Structure of Scientific Revolutions* (The University of Chicago Press).
- Lecun, Y., Bottou, L., Bengio, Y., et al. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324.
- Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90.
- Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. *J. Mach. Learn. Res.* **15**, 315–323.
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. Preprint at arXiv.
- Lin, M., Chen, Q., and Yan, S. (2013). Network in network. Preprint at arXiv.
- Szegedy, C., Liu, W., Jia, Y., et al. (2014). Going deeper with convolutions. Preprint at arXiv.
- He, K.M., Zhang, X.Y., Ren, S.Q., et al. (2016). Deep residual learning for image recognition. Preprint at arXiv.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2020). An image is worth 16x16 words: transformers for image recognition at scale. Preprint at arXiv.
- Liu, Z., Lin, Y., Cao, Y., et al. (2021). Swin transformer: hierarchical vision transformer using shifted windows. Preprint at arXiv.
- Rombach, R., Blattmann, A., Lorenz, D., et al. (2022). High-resolution image synthesis with latent diffusion models. Preprint at arXiv.
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial nets. Preprint at arXiv.
- Kirillov, A., Mintun, E., Ravi, M., et al. (2023). Segment anything. Preprint at arXiv.
- Cheng, C.-C., Hwu, Y., and Ching, Y.-T. (2016). High-resolution and large-volume tomography reconstruction for X-ray microscopy. *SPIE Med. Imaging (Portsm. R.I.)*.
- Vescovi, R., Du, M., de Andrade, V., et al. (2018). Tomosaic: efficient acquisition and reconstruction of teravoxel tomography data using limited-size synchrotron X-ray beams. *J. Synchrotron Radiat.* **25**, 1478–1489.
- Qin, G., Yang, C., and Du, Y. (2021). Image enhancement of optical coherence tomography using deep learning. *AIID*.
- Li, M., Huang, K., Xu, Q., et al. (2012). OCTA-500: a retinal dataset for optical coherence tomography angiography study. Preprint at arXiv.
- Guizar-Sicairos, M., Diaz, A., Holler, M., et al. (2011). Phase tomography from x-ray coherent diffractive imaging projections. *Opt Express* **19**, 21345–21357.
- Liu, Y., Meirer, F., Williams, P.A., et al. (2012). TXM-Wizard: a program for advanced data collection and evaluation in full-field transmission X-ray microscopy. *J. Synchrotron Radiat.* **19**, 281–287.
- Han, R., Zhang, F., Wan, X., et al. (2014). A marker-free automatic alignment method based on scale-invariant features. *J. Struct. Biol.* **186**, 167–180.
- Yu, H., Xia, S., Wei, C., et al. (2018). Automatic projection image registration for nanoscale X-ray tomographic reconstruction. *J. Synchrotron Radiat.* **25**, 1819–1826.
- Zhang, Z., Bi, X., Li, P., et al. (2023). Automatic synchrotron tomographic alignment schemes based on genetic algorithms and human-in-the-loop software. *J. Synchrotron Radiat.* **30**, 169–178.
- Gürsoy, D., Hong, Y.P., He, K., et al. (2017). Rapid alignment of nanotomography data using joint iterative reconstruction and reprojection. *Sci. Rep.* **7**, 11818.
- Pande, K., Donatelli, J.J., Parkinson, D.Y., et al. (2022). Joint iterative reconstruction and 3D rigid alignment for X-ray tomography. *Opt Express* **30**, 8898–8916.
- Zhang, Z., Dong, Z., Yan, H., et al. (2023). A general image misalignment correction method for tomography experiments. *iScience* **26**, 107932.
- Yang, X., De Carlo, F., Phatak, C., et al. (2017). A convolutional neural network approach to calibrating the rotation axis for X-ray computed tomography. *J. Synchrotron Radiat.* **24**, 469–475.
- Topal, E., Löffler, M., and Zschech, E. (2020). Deep learning-based inaccuracy compensation in reconstruction of high resolution XCT data. *Sci. Rep.* **10**, 7682.
- Fu, T., Zhang, K., Wang, Y., et al. (2022). Feature detection network-based correction method for accurate nano-tomography reconstruction. *Appl. Opt.* **61**, 5695–5703.
- Fu, T., Zhang, K., Wang, Y., et al. (2021). Deep-learning-based image registration for nano-resolution tomographic reconstruction. *J. Synchrotron Radiat.* **28**, 1909–1915.
- Liu, M., Han, Y., Xi, X., et al. (2022). Multiscale dense U-Net: a fast correction method for thermal drift artifacts in laboratory NanoCT scans of semi-conductor chips. *Entropy* **24**, 967.
- Cho, S.-J., Ji, S.-W., Hong, J.-P., et al. (2021). Rethinking coarse-to-fine approach in single image deblurring. Preprint at arXiv.
- Huang, G., Liu, Z., van der Maaten, L., et al. (2017). Densely connected convolutional networks. *CVPR*.
- Oord, A.V.D., Li, Y., and Vinyals, O. (2018). Representation learning with contrastive predictive coding. Preprint at arXiv.
- He, K., Fan, H., Wu, Y., et al. (2020). Momentum contrast for unsupervised visual representation learning. *CVPR*.
- Chen, X., Fan, H., Girshick, R., et al. (2020). Improved baselines with momentum contrastive learning. Preprint at arXiv.
- Prannay, K., Teterwak, P., Wang, C., et al. (2020). Supervised contrastive learning. Preprint at arXiv.

66. Chen, T., Kornblith, S., Norouzi, M., et al. (2020). A simple framework for contrastive learning of visual representations. Preprint at arXiv.
67. Liang, K., Zhang, L., Yang, Y., et al. (2018). A self-supervised deep learning network for low-dose CT reconstruction. *NSS/MIC*.
68. Lee, H., Lee, J., Kim, H., et al. (2019). Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Trans. Radiat. Plasma Med. Sci.* **3**, 109–119.
69. Okamoto, T., Ohnishi, T., and Haneishi, H. (2022). Artifact reduction for sparse-view CT using deep learning with band patch. *IEEE Trans. Radiat. Plasma Med. Sci.* **6**, 859–873.
70. Jin, K.H., McCann, M.T., Froustey, E., et al. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522.
71. Okamoto, T., Kumakiri, T., and Haneishi, H. (2022). Patch-based artifact reduction for three-dimensional volume projection data of sparse-view micro-computed tomography. *Radiol. Phys. Technol.* **15**, 206–223.
72. Xie, S., Zheng, X., Chen, Y., et al. (2018). Artifact removal using improved GoogLeNet for sparse-view CT reconstruction. *Sci. Rep.* **8**, 6700.
73. Lehtinen, J., Munkberg, J., Hasselgren, J., et al. (2018). Noise2Noise: learning image restoration without clean data. Preprint at arXiv.
74. Gao, X., Su, T., Zhang, Y., et al. (2023). Attention-based dual-branch deep network for sparse-view computed tomography image reconstruction. *Quant. Imag. Med. Surg.* **13**, 1360–1374.
75. Raparia, D., Alessi, J., and Kponou, A. (1997). The algebraic reconstruction technique (ART). Preprint at arXiv.
76. Bappy, D.M., and Jeon, I. (2017). Modified simultaneous iterative reconstruction technique for fast, high-quality CT reconstruction. *IET Image Process.* **11**, 701–708.
77. Huang, Y., Wang, S., Guan, Y., et al. (2020). Limited angle tomography for transmission X-ray microscopy using deep learning. *J. Synchrotron Radiat.* **27**, 477–485.
78. Wang, J., Liang, J., Cheng, J., et al. (2020). Deep learning based image reconstruction algorithm for limited-angle translational computed tomography. *PLoS One* **15**, e0226963.
79. Yiran, J., McMichael, N., Mokarzel, P., et al. (2020). Algorithm for limited angle CT reconstruction with U-net based regularization. *NSS/MIC*.
80. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* **521**, 436–444.
81. Yang, X., Kahnt, M., Brückner, D., et al. (2020). Tomographic reconstruction with a generative adversarial network. *J. Synchrotron Radiat.* **27**, 486–493.
82. Wang, C., Ding, G., Liu, Y., and Xin, H.L. (2020). 0.7 angstrom resolution electron tomography enabled by deep-learning-aided information recovery. *Adv. Intell. Syst.* **2**, 2000152.
83. Xin, H., Wei, H., Ding, G., et al. (2020). Artificial intelligence enabled information inpainting and artifact removal for electron tomography. *Microsc. Microanal.* **26**, 664–665.
84. Barutcu, S., Aslan, S., Katsaggelos, A.K., et al. (2021). Limited-angle computed tomography with deep image and physics priors. *Sci. Rep.* **11**, 17740.
85. Lempitsky, V., Vedaldi, A., and Ulyanov, D. (2018). Deep image prior. *CVPR*.
86. Qiu, B., You, Y., Huang, Z., et al. (2021). N2NSR-OCT: simultaneous denoising and super-resolution in optical coherence tomography images using semisupervised deep learning. *J. Biophotonics* **14**, e202000282.
87. Jing, J., Xia, W., Hou, M., et al. (2022). Training low dose CT denoising network without high quality reference data. *Phys. Med. Biol.* **67**, 084002.
88. Chen, H., Zhang, Y., Kalra, M.K., et al. (2017). Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans. Med. Imaging* **36**, 2524–2535.
89. Selim, M., Zhang, J., Brooks, M.A., et al. (2023). DiffusionCT: latent diffusion model for CT image standardization. Preprint at arXiv.
90. Liu, Z., Bicer, T., Kettimuthu, R., et al. (2020). TomoGAN: low-dose synchrotron x-ray tomography with generative adversarial networks: discussion. *J. Opt. Soc. Am. Opt. Image Sci. Vis.* **37**, 422–434.
91. Shan, H., Zhang, Y., Yang, Q., et al. (2018). 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *IEEE Trans. Med. Imaging* **37**, 1522–1534.
92. Yang, L., Li, Z., Ge, R., et al. (2023). Low-dose CT denoising via sinogram inner-structure transformer. *IEEE Trans. Med. Imaging* **42**, 910–921.
93. Yin, X., Zhao, Q., Liu, J., et al. (2019). Domain progressive 3D residual convolution network to improve low-dose CT imaging. *IEEE Trans. Med. Imaging* **38**, 2903–2913.
94. Zhou, Z., Li, C., Bi, X., et al. (2023). A machine learning model for textured X-ray scattering and diffraction image denoising. *NPJ Comput. Mater.* **9**, 58.
95. Hendriksen, A.A., Pelt, D.M., and Batenburg, K.J. (2020). Noise2Inverse: self-supervised deep convolutional denoising for tomography. *IEEE Trans. Comput. Imaging* **6**, 1320–1335.
96. De Carlo, F., Gürsöy, D., Ching, D.J., et al. (2018). TomoBank: a tomographic data repository for computational x-ray science. *Meas. Sci. Technol.* **29**, 034004.
97. Yang, H., Wang, W., Shang, J., et al. (2021). Segmentation of computed tomography images and high-precision reconstruction of rubber composite structure based on deep learning. *Compos. Sci. Technol.* **213**, 108875.
98. Davydzhenka, T., Sinclair, D., Chawla, N., et al. (2022). Deep-layers-assisted machine learning for accurate image segmentation of complex materials. *Mater. Charact.* **192**, 112175.
99. Torbati-Sarrafi, H., Niverty, S., Singh, R., et al. (2021). Machine-learning-based algorithms for automated image segmentation techniques of transmission X-ray microscopy (TXM). *Jom* **73**, 2173–2184.
100. Lin, B., Emami, N., Santos, D.A., et al. (2022). A deep learned nanowire segmentation model using synthetic data augmentation. *NPJ Comput. Mater.* **8**, 88.
101. Zhou, Z.W., Siddiquee, M.M.R., Tajbakhsh, N., et al. (2018). UNet plus plus : a nested U-Net architecture for medical image segmentation. Preprint at arXiv.
102. Zhao, H.S., Shi, J.P., Qi, X.J., et al. (2017). Pyramid scene parsing network. Preprint at arXiv.
103. Chen, L.-C., Zhu, Y., Papandreou, G., et al. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. Preprint at arXiv.
104. Deng, J., Dong, W., Socher, R., et al. (2009). ImageNet: a large-scale hierarchical image database. *CVPR* **1**.
105. He, K., Gkioxari, G., Dollár, P., et al. (2020). Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 386–397.
106. Moebel, E., Martinez-Sanchez, A., Lamm, L., et al. (2021). Deep learning improves macromolecule identification in 3D cellular cryo-electron tomograms. *Nat. Methods* **18**, 1386–1394.
107. Liu, C., Zeng, X., Lin, R., et al. (2018). Deep learning based supervised semantic segmentation of Electron Cryo-Subtomograms. *ICIP* **2018**, 1578–1582.
108. Zhu, H., Wang, C., Wang, Y., et al. (2022). Unsupervised multi-task learning for 3D subtomogram image alignment, Clustering and Segmentation. *ICIP*.
109. Zhou, B., Yu, H., Zeng, X., et al. (2020). One-shot learning with attention-guided segmentation in cryo-electron tomography. *Front. Mol. Biosci.* **7**, 613347.
110. Heebner, J.E., Purnell, C., Hylton, R.K., et al. (2022). Deep learning-based segmentation of cryo-electron tomograms. *J. Vis. Exp.*
111. Wu, S., Liu, G., and Yang, G. (2022). Fast particle picking for cryo-electron tomography using one-stage detection. *ISBI*.
112. Genc, A., Kovarik, L., and Fraser, H.L. (2022). A deep learning approach for semantic segmentation of unbalanced data in electron tomography of catalytic materials. *Sci. Rep.* **12**, 16267.
113. Nguyen, N., Bohak, C., Engel, D., et al. (2023). Finding Nano-Otzi: cryo-electron tomography visualization guided by learned segmentation. *IEEE Trans. Vis. Comput. Graph.* **29**, 4198–4214.
114. Zhang, S., Liu, J., Hu, B., et al. (2023). GH-DDM: the generalized hybrid denoising diffusion model for medical image generation. *Multimed. Syst.* **29**, 1335–1345.
115. Wu, J., Fu, R., Fang, H., et al. (2022). MedSegDiff: medical image segmentation with diffusion probabilistic model. Preprint at arXiv.
116. Wu, J., Fu, R., Fang, H., et al. (2023). MedSegDiff-V2: diffusion based medical image segmentation with transformer. Preprint at arXiv.
117. Kopp, R., Joseph, J., Ni, X., et al. (2022). Deep learning unlocks X-ray microtomography segmentation of multiclass microdamage in heterogeneous materials. *Adv. Mater.* **34**, 2107817.
118. Fu, T., Monaco, F., Li, J., et al. (2022). Deep-learning-enabled crack detection and analysis in commercial lithium-ion battery cathodes. *Adv. Funct. Mater.* **32**, 2203070.
119. Shorten, C., and Khoshgoftaar, T.M. (2019). A survey on image data augmentation for deep learning. *JBD* **6**, 60.
120. Thersleff, T., and Tai, C.-W. (2023). Feature-specific correlation of structural, optical, and chemical properties in the transmission electron microscope with hypermodal data fusion. *Microsc. Microanal.* **29**, 166–179.
121. Muto, S., and Shiga, M. (2020). Application of machine learning techniques to electron microscopic/spectroscopic image data analysis. *Microscopy* **69**, 110–122.
122. Nandhini Abirami, R., Durai Raj Vincent, P.M., Srinivasan, K., et al. (2022). Multimodal medical image fusion of positron emission tomography and magnetic resonance imaging using generative adversarial networks. *Behav. Neuro.* **2022**, 6878783.
123. Liu, M., and Mukerji, T. (2022). Multiscale fusion of digital rock images based on deep generative adversarial networks. *Geophys. Res. Lett.* **49**.
124. Karras, T., Laine, S., Aittala, M., et al. (2020). Analyzing and improving the image quality of StyleGAN. *CVPR*.
125. McKenzie, E.M., Santhanam, A., Ruan, D., et al. (2020). Multimodality image registration in the head-and-neck using a deep learning-derived synthetic CT as a bridge. *Med. Phys.* **47**, 1094–1104.
126. Blendowski, M., Bouteldja, N., and Heinrich, M.P. (2020). Multimodal 3D medical image registration guided by shape encoder-decoder networks. *IJCARS* **15**, 269–276.
127. Hering, A., Kuckertz, S., Heldmann, S., et al. (2019). Memory-efficient 2.5D convolutional transformer networks for multi-modal deformable registration with weak label supervision applied to whole-heart CT and MRI scans. *IJCARS* **14**, 1901–1912.
128. Badae, C.T., Holbrook, M., and Clark, D.P. (2018). Multi-energy CT decomposition using convolutional neural networks. *Med. Imaging*.
129. Gong, H., Tao, S., Rajendran, K., et al. (2020). Deep-learning-based direct inversion for material decomposition. *Med. Phys.* **47**, 6294–6309.
130. Nadkarni, R., Allphin, A., Clark, D.P., et al. (2022). Material decomposition from photon-counting CT using a convolutional neural network and energy-integrating CT training labels. *Phys. Med. Biol.* **67**, 155003.
131. Fang, W., Wu, D., Kim, K., et al. (2020). Direct dual energy CT material decomposition using Noise2Noise prior. *NSS/MIC*.
132. Terzi, S., Salvo, L., Suéry, M., et al. (2009). In situ X-ray tomography observation of inhomogeneous deformation in semi-solid aluminium alloys. *Scr. Mater.* **61**, 449–452.

133. Yang, Z., Ren, W., Sharma, R., et al. (2017). In-situ X-ray computed tomography characterisation of 3D fracture evolution and image-based numerical homogenisation of concrete. *Cem. Concr. Compos.* **75**, 74–83.
134. Scott, A.E., Mavrogordato, M., Wright, P., et al. (2011). In situ fibre fracture measurement in carbon-epoxy laminates using high resolution computed tomography. *Compos. Sci. Technol.* **71**, 1471–1477.
135. Lorenzoni, R., Curosu, I., Léonard, F., et al. (2020). Combined mechanical and 3D-microstructural analysis of strain-hardening cement-based composites (SHCC) by in-situ X-ray microtomography. *Cem. Concr. Res.* **136**, 106139.
136. Wang, T., Jiang, Z., Kemao, Q., et al. (2016). GPU accelerated digital volume correlation. *Exp. Mech.* **56**, 297–309.
137. Duan, X., and Huang, J. (2022). Deep learning-based digital volume correlation. *Extreme Mech. Lett.* **53**, 101710.
138. Lialin, V., Deshpande, V., and Rumshisky, A. (2023). Scaling down to scale up: a guide to parameter-efficient fine-tuning. Preprint at arXiv.
139. Chen, Z., Duan, Y., Wang, W., et al. (2023). Vision transformer adapter for dense predictions. Preprint at arXiv.
140. Yang, T., Zhu, Y., Xie, Y., et al. (2023). AIM: adapting image models for efficient video action recognition.
141. Zhou, K., Yang, J., Loy, C.C., et al. (2022). Learning to prompt for vision-language models. *IJCV* **130**, 2337–2348.
142. Jia, M.L., Tang, L.M., Chen, B.C., et al. (2022). Visual prompt tuning. Preprint at arXiv.
143. Vijayakumar, J., Yuan, H., Mille, N., et al. (2023). Soft X-ray spectro-ptychography of boron nitride nanobamboos, carbon nanotubes and permalloy nanorods. *J. Synchrotron Radiat.* **30**, 746–757.
144. Radford, A., Kim, J.W., Hallacy, C., et al. (2021). Learning transferable visual models from natural language supervision. Preprint at arXiv.
145. Bai, Y., Jones, A., Ndousse, K., et al. (2022). Training a helpful and harmless assistant with reinforcement learning from human feedback. Preprint at arXiv.

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation for Young Scientists of China (grant 12005253), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB 37000000), the Innovation Program of the Institute of High Energy Physics, CAS (E25455U210), and the Hefei Science Center, Chinese Academy of Sciences (award 2019HSC-KPRD003). All authors gratefully acknowledge support from the BL13HB and BL16U2 beamlines of the Shanghai Synchrotron Radiation Facility (SSRF) and BL07W beamline of the National Synchrotron Radiation Laboratory (NSRL).

AUTHOR CONTRIBUTIONS

Y.Z. and Z.D. conceived the project. Z.Z. and C.L. wrote the manuscript. W.W., G.L., and Y.D. revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

LEAD CONTACT WEBSITE

<https://people.ucas.edu.cn/~zhangyigxj/>