OPTICA

Deep learning enhanced super-resolution x-ray fluorescence microscopy by a dual-branch network

Xiaoyin Zheng,[1] Varun R. Kankanallu,[1] Chang-An Lo,[1] Ajith Pattammattel,[2] Yong Chu,[2] Yu-Chen Karen Chen-Wiegart,[1,2,3] and Xiaojing Huang[2,*]

[1]Department of Materials Science and Chemical Engineering, Stony Brook University, Stony Brook, New York 11794, USA
[2]Brookhaven National Laboratory, National Synchrotron Light Source II, Upton, New York 11973, USA
[3]Karen.Chen-Wiegart@stonybrook.edu
*xjhuang@bnl.gov

Received 12 September 2023; revised 26 December 2023; accepted 28 December 2023; published 25 January 2024

X-ray fluorescence (XRF) microscopy is a powerful technique for quantifying the distribution of elements in complex materials, which makes it a crucial imaging technique across a wide range of disciplines in physical and biological sciences, including chemistry, materials science, microbiology, and geosciences. However, as a scanning microscopy technique, the spatial resolution of XRF imaging is inherently constrained by the x-ray probe profile and scanning step size. Here we propose a dual-branch machine learning (ML) model, which can extract scale-variant features and bypass abundant low-frequency information separately, to enhance the spatial resolution of the XRF images by mitigating the effects of blurring from the probe profile. The model is trained by simulated natural images, and a two-stage training strategy is used to overcome the domain gap between the natural images and experimental data. The tomography reconstruction from enhanced XRF projections shows an improvement in resolution by a scale factor of four and reveals distinct internal features invisible in low-resolution XRF within a battery sample. This study offers a promising approach for obtaining high-resolution XRF imaging from its low-resolution version, paving the way for future investigations in a broader range of disciplines and materials.

https://doi.org/10.1364/OPTICA.503398

1. INTRODUCTION

Hard x-ray fluorescence (XRF) microscopy offers unique capabilities in probing elemental distributions in advanced materials, experiencing a marked increase in its application across diverse material science domains [1]. This technique facilitates the concurrent mapping of multiple elements, resulting in intricate elemental colocalization maps. A notable strength of XRF lies in its ability to quantitatively analyze elemental distribution within materials, shedding light on the structural and compositional nuances of materials. The use of x-ray excitation significantly reduces the Bremsstrahlung background compared with the background from high-energy electrons [2], amplifying the elemental sensitivity of hard XRF microscopy. Moreover, the high penetration depth of x-rays, extending up to tens of microns, positions XRF microscopy as a premier choice for tomographic visualization of material samples.

In addition, multimodal imaging, which combines XRF microscopy with other imaging modalities, can provide complementary scientific information. Facilities like the Hard X-ray Nanoprobe (HXN) beamline at the National Synchrotron Light Source II (NSLS-II) , Brookhaven National Laboratory [3,4], enable the concurrent collection of XRF and ptychography data in a fly-scan [5] mode. The same technique is also used in the I12 beamline of the Diamond Light Source (DLS) [6] and the Advanced Photon Source (APS) at Argonne National Laboratory [7]. In this mode, the sample remains in continuous motion, while the detector synchronously records data, thus eliminating the majority of the motion overhead. Fly-scan boosts the throughput of ptychography and XRF imaging, especially for three-dimensional applications and dynamic systems. However, the XRF images measured in a fly-scan mode can be further blurred by the movement of the sample and probe. Thus, it becomes challenging to resolve fine features within the samples, which hinders the understanding of the detailed morphological and elemental distributions that often critically determine their functionalities.

Enhancement of the XRF images to reveal detailed features is feasible as the XRF intensity measurement can be modeled as a convolution between the sample and probe integrated with impacts from the scanning step sizes and the fly-scan mode. Including multiple illumination modes into the ptychography reconstruction engines has been proven to effectively decouple the impact from continuous motion and recover both the probe modes and the sample function [8]. As we know the scanning step size and how the fly-scan is executed, together with the probe function recovered from the ptychography reconstruction, we can in principle generate the equivalent kernel in the convolution process of XRF measurement, which makes it possible to conduct a deconvolution and enhance the resolution of XRF images [9,10].

Recently, the development of single image super resolution by neural networks has shown significant advantage over traditional methods and provides a valid solution for this problem [11,12]. Inspired from the sparse-coding-based method [13], one of the representative external example-based traditional super resolution methods relying on building a mapping between the low-resolution dictionary encoded by cropped patches with overlapping and high-resolution dictionary. Dong *et al.* [12] designed the super-resolution convolution neural network (SRCNN) and achieved excellent performance over the traditional super resolution methods [14]. Subsequently, advancements in deep learning architectures lead to the development of deeper and wider networks tailored for super-resolution tasks, such as enhanced deep super-resolution (EDSR) [15] and the multi-scale deep super-resolution (MDSR) [16]. EDSR focuses on optimizing performance by removing unnecessary modules in conventional residual networks, while MDSR is designed to handle multiple scale factors within a single model, showcasing the versatility and efficiency of these architectures.

Wu *et al*. [17] demonstrated enhancing the spatial resolution of XRF microscopy using machine learning techniques, specifically addressing the limitations imposed by the x-ray probe size. This approach is particularly relevant to our work, as it demonstrates the feasibility of machine learning models in overcoming physical constraints in imaging methods, and successfully decoupled the impact of the x-ray probe from the XRF signal using an RDB-based network. Inspired by the prior work, here we introduce a dual-branch model that combines both the residual channel attention network (RCAN) [18] and the multi-scale residual network (MSRN) [19]. The channel attention mechanism used in RCAN adaptively rescales channel-wise features by considering interdependencies among channels, thereby enhancing the representational ability of convolutional neural networks. This leads to improved accuracy and visual enhancements in image super-resolution tasks, outperforming the state-of-the-art methods. The multi-scale residual blocks (MSRB) used in MSRN utilize convolution kernels of varying sizes to adaptively detect image features at different scales, allowing these features to interact and extract the most effective image information. In the work demonstrated by Wu *et al.* [17], the fly-scan on natural images was incorporated into the training and successfully enhanced spatial resolution in experimental data. However, in deep learning, the domain gap between XRF data and natural images is a persistent issue, particularly due to the distinct noise characteristics and inherent complexities of XRF datasets. These XRF datasets often exhibit unique noise patterns, such as photon-counting noise or background fluorescence, which are markedly different from the types of noise encountered in natural images. This domain gap requires specialized adaptation strategies for models initially trained on the natural image datasets, where noise typically manifests as pixilation, compression artifacts, or variations in lighting and color. Thus, we employ a two-stage training strategy: the network was trained using natural image datasets from the DIV2K dataset [20], followed by fine-tuning with experimental data from the HXN beamline of NSLS-II. This strategy addresses the limited availability of experimental data and the domain gap between experimental data and natural images. Each projection was enhanced individually, and tomographic reconstruction was performed to visualize the 3D structure of a charged Zn/MnO_2 battery cathode in an aqueous electrolyte for potential grid-scale energy storage applications [15]. Experimentally, the data acquisition in a fly-scan mode can effectively enhance the scanning rate by continuously moving the scanning stages, minimizing overhead. However, the resulting images are a moving average of the convolution between the beam profile and the sample. Here we incorporated the fly-scan process in the training of the neutral network model. Our method successfully mitigated the blurring effect caused by the scanning process and restored the fine structure of cracks and pores present in the sample. This method provides a reliable means of obtaining high-resolution elemental distribution and enables more accurate morphological analysis, particularly for nanoscale structures.

In summary, the novelty of the work is as follows. We introduced a novel dual-branch network architecture that combines residual channel attention blocks (RCAB) and multi-scale residual blocks (MSRB). The network focuses on extracting features across various spatial scales and uniquely prioritizes high-frequency feature channels, enabling XRF super resolution. We adopted a two-stage training process, initially training the network on simulated natural image pairs blurred in fly-scan mode, followed by refinement with experimental data to suppress noise. This method, coupled with enhancements in both angular and lateral dimensions, achieved a fourfold improvement in resolution. The effectiveness of our fine-tuned network was further validated by its capacity to produce high-quality image reconstructions from limited projection data, achieving image quality comparable to that from full projections.

2. METHODS

A. Simulation of X-Ray Fluorescence Microscopy in Fly-Scan Mode

In the fly-scan mode, the convolution between sample and beam profile can be described as [17]

$$Y_j(Z, \boldsymbol{r}) = \sigma(Z, \lambda) \int_{\text{to}}^{t0+\Delta t} \int_{-\infty}^{+\infty} P(\boldsymbol{r} - \boldsymbol{v}t) N(Z, \boldsymbol{r} + \boldsymbol{r}_j) \mathrm{d}r \mathrm{d}t, \tag{1}$$

where $\boldsymbol{r}$ represents two-dimensional coordinates in real space, $\boldsymbol{v}$ is the scan speed, Z is the atomic number, and λ is the wavelength of incident x-ray. $Y_j(Z, \boldsymbol{r})$ is the measured XRF and $\sigma(Z, \lambda)$ is XRF cross section [21]. $P(\boldsymbol{r})$ is the beam profile, and $N(Z, \boldsymbol{r})$ is the elemental distribution of the measures sample. It is evident that the observed elemental distribution is affected by spatial blur due to the moving beam profile and temporal blur from the scanning duration. Given that the probe profile can be derived from ptychographic reconstruction, it is feasible to differentiate the double integration and estimate the actual measured sample accurately. This enhancement improves the resolution of XRF, providing a more precise elemental distribution for quantitative data analysis.

To simulate the fly-scan [5] process, a discrete version of Eq. (1) was utilized on natural images from the DIV2K dataset [20] to establish training pairs. As depicted in Fig. 1(a), the probe advanced in the scanning direction. At each scanning spot, the sum of the element-wise product between the probe profile and the scanning region in the high-resolution image was computed. The summation across all scanning spots was then averaged to determine the intensity of the corresponding pixel in the low-resolution images. In the simulation, the pixel size of the probe is 7.5 nm, and the scanning step size is every 4 pixels, which is determined by the ratio of the raw XRF pixel size (30 nm) to the pixel size of the phase-contrast image from ptychography reconstruction (7.5 nm). In Fig. 1(b), the ground truth image represents

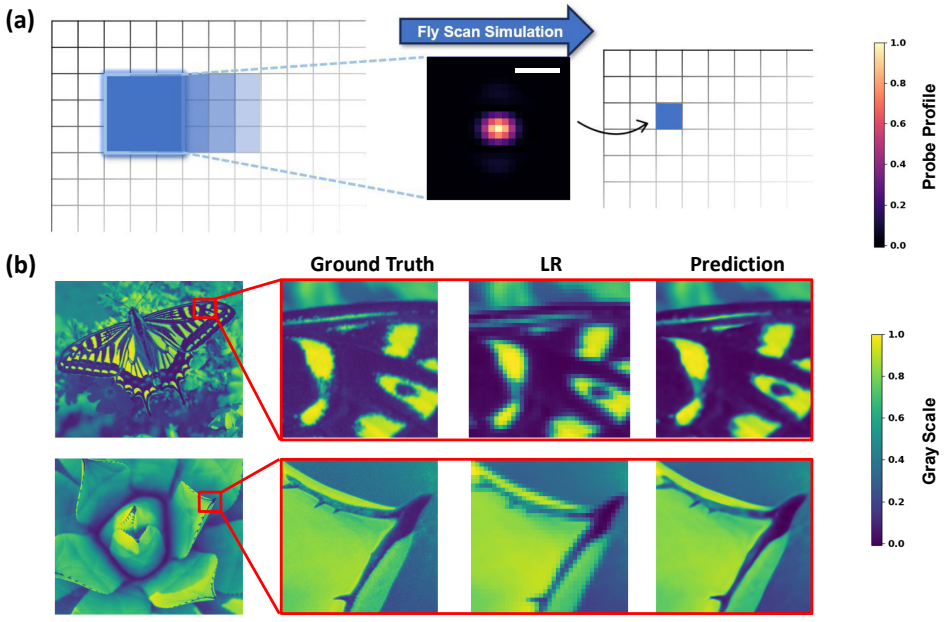

Fig. 1. Schematic of the fly-scan simulation and the enhancement results (prediction) on natural images by the proposed network. (a) demonstrates the simulated fly-scan process. In this simulation, the probe profile, obtained from ptychography reconstruction, scans a high-resolution image using a specific step size. At every step of the scanning process, the element-wise product of the probe profile and the scanned area is calculated and then averaged on the training dataset to form image pairs. These pairs help the model learn and reduce the blurring effects caused by the probe profile. The scale bar is 50 nm, and the pixel size of the probe is 7.5 nm. (b) shows an enlarged view of both the ground truth and the low-resolution image from the validation dataset. The low-resolution image is enhanced by a factor of four, resulting in a prediction that aligns with the size of the ground truth image. It can be observed that the model successfully recaptures detailed features in the image.

the measured sample, while the low-resolution image from the fly-scan corresponds to the elemental distribution. Notably, the feature boundaries in the low-resolution image appear considerably blurred due to scanning with the probe profile. Additionally, the high-resolution image was down-sampled by the scanning step size, resulting in pixel-like features.

B. Machine Learning Model for Super Resolution

Figure 2 shows the structure of the proposed model inspired from RCAB [18] and MSRB [19]. Super resolution of 4 times is considered in this case, i.e., the width and height of the high-resolution output image is 4 times larger than the low-resolution input image. The proposed model has a dual-branch architecture to extract features at different space scale and frequency scale. Four fundamental modules are here to form the complete proposed model for XRF image super resolution: (i) a shallow feature extraction module to extract low level features from the input, (ii) a spatial extraction module to extract features at different scale (see Fig. S1 in Supplement 1), (iii) a frequency filter module to bypass the abundant low frequency information and focus on the high frequency information representing detailed structure (see Fig. S2 in Supplement 1), and (iv) an up-sampling module to for upscaling the final high-resolution output from low resolution feature maps.

The spatial feature extraction module consists of several multi scale residual blocks (MSRB). In each MSRB, there is a two-bypass structure where each bypass utilizes a unique kernel size. The feature maps extracted from each bypass are then concatenated to share the information from different spatial scales. A local residual learning is also employed to enhance the learning efficiency. In the frequency filter module consisting of residual channel attention blocks (RCAB), a combination of the short and long skip connection is used to filter the low frequency information representing the texture and the background. Low-frequency information and high-frequency information are of different importance in recovering high-resolution images. Treating them equally across channels limits the representational ability of convolutional neural networks (CNNs). To overcome this issue, a channel attention mechanism [22,23] is used in RCAB to adaptively rescale channel-wise features by considering interdependencies among channels. By leveraging channel interdependencies, the mechanism refines feature maps, ensuring that more relevant channels are amplified while less pertinent ones are suppressed, thereby optimizing feature representation for improved super-resolution performance. A residual in residual (RIR) structure, composed of residual groups (RG), is introduced, forming a very deep network composed of several residual groups with long skip connections. Each of these RG contains residual blocks with short skip connections. This channel attention mechanism, combined with the RIR structure, allows the network to bypass the low-frequency information and restore the high-resolution image more accurately. Contrary to traditional methods that upscale the low-resolution (LR) input image to high-resolution (HR) before reconstruction [24,25], adding computational complexity, the used up-sampling layer extracts feature maps in the LR space. The sub-pixel convolution layer [26] then upscales these final LR feature maps directly to the HR output. This method replaces the conventional bicubic filter with upscaling filters specifically trained for each feature map, offering enhanced performance and swifter processing compared to prior CNN-based techniques.

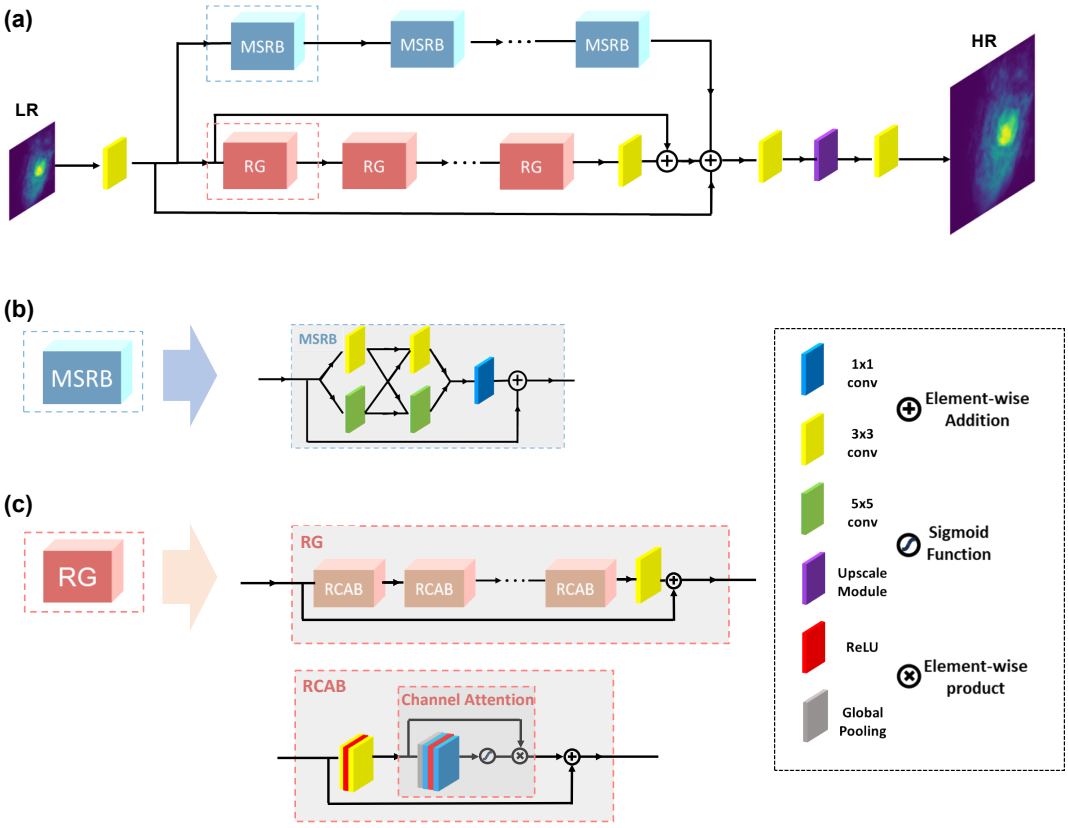

Fig. 2. (a) Architecture of the proposed dual-branch model. Multi-scale residual block (MSRB) and residual channel attention blocks (RCAB) serve as the foundational modules, each tailored to extract features from distinct domains: the former focuses on the spatial domain while the latter targets the frequency domain. Subsequent to their individual operations, the feature channels from both branches are added element-wise. This combined channel is then fed into a pixel shuffle layer, which up-samples the image to the desired resolution. The scale factor employed in this architecture is 4. (b) presents the architecture of the MSRB, which consists of convolutional kernels of two different sizes. (c) displays the residual group (RG) composed of RCABs. In RCAB, the channel attention mechanism is employed to selectively bypass the channels with low-frequency features.

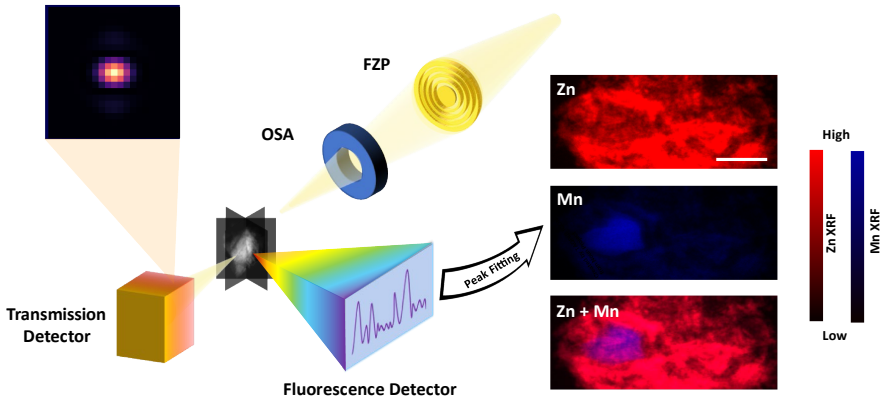

Fig. 3. Schematic illustration of the experimental setup at the Hard X-ray Nanoprobe (HXN) beamline of NSLS-II [3]. The incoming x-ray nanobeam is focused by a Fresnel zone plate (FZP). An order-sorting-aperture (OSA) following the FZP removes unwanted background to ensure a clear focus. The sample is rotated and scanned in fly-scan mode at various angles. A fluorescence detector captures signals for elemental mapping, while a transmission detector is used for phase-contrast imaging. The XRF mapping is fitted with the Mn and Zn emission lines. The scale bar is 2 μm.

3. EXPERIMENTS

A. Experimental Setup

A charged zinc/manganese oxide (Zn/MnO_2) battery cathode in a mild aqueous electrolyte was studied after 32 cycles (450 h) at a C-rate of 0.1C. The sample was prepared following the same procedure as described in our previous work [27]. The cycled cathode was extracted in its charge state followed by immediate washing in HPLC grade water and then was dried under vacuum for 12 h. The electrode was then scrapped from the current collector directly onto a sample mounting diving board. The experimental setup is shown in Fig. 3. During data acquisition at the Hard X-ray Nanoprobe (HXN) beamline, the x-ray nanobeam was focused by a Fresnel zone plate (FZP), and the sample was measured in

fly-scan mode at 10 keV. The zone plate used in this measurement has 30 nm outermost zone width and 250 μm diameter. The focus size is 40 nm. The dwell time per scan point was 0.03 s. Two locations of the same sample were measured. One was measured at 30 nm and 120 nm scanning step size, with the projections collected every 2° from 90° to −90° and was used to demonstrate the super-resolution imaging enhancement. The other was utilized as the fine-tuning data and measured at 50 nm and 200 nm scanning step size, with the projections collected every 3° from 90° to −18°. The fluorescence signal was collected by an energy-dispersive detector, and the diffraction pattern was collected by a pixel-array detector. The probe profile used for simulation was obtained by ptychography reconstruction from the 90° projection at 30 nm. The goal is to identify the irreversible Zn Mn complex [28], along with the morphological changes.

B. Implementation Details and Training

In the proposed spatial extraction module of the network, which utilizes a two-bypass structure, filters of sizes 3×3 and 5×5 are employed to detect features across different scales. These filters are followed by a 1×1 convolutional layer, specifically designed to amalgamate the local features. In the frequency filter module, which comprises RCABs, all convolutional layers predominantly use a 3×3 filter size. However, the channel attention employs a kernel size of 1×1. The network first compresses the number of feature layers to one-sixteenth of its original size, then expands it by a factor of 16. The top branch integrates five MSRBs. The bottom branch includes 10 RGs, with each RG containing 10 RCABs. The network achieves an overall up-sampling factor of 4 through two pixel-shuffle layers, each capable of up-sampling by a factor of 2. The activation function employed is the rectified linear unit (ReLU) [29]. For backpropagation, the ADAM [30] optimizer is chosen with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The training process optimizes the L1 loss function [31]. The learning rate starts at 10^{-4} and is decreased to one half after the first 100 epochs. Subsequently, it decays by half every 50 epochs, with the model being trained for a total of 250 epochs. The implementation of the model was executed in Python using PyTorch [32], and the training was conducted on an NVIDIA 3090Ti GPU. It took 19 h to train the simulated natural images and 4 h to fine-tune the network by experimental data. In our experiment, obtaining 91 projections with a pixel size of 30 nm took 16 h. In contrast, acquiring the low-resolution data (a coarser scanning step size), with a pixel size of 120 nm, only required approximately 1 h. Without sacrificing achievable resolution, this enhancement in data acquisition throughput and dose efficiency would make a huge impact on scientific applications, especially for radiation-sensitive sample systems.

The model was trained using natural images from the DIV2K dataset [20]. To generate the low-resolution image, the high-resolution image underwent a fly-scan using the probe profile derived from ptychography reconstruction, with a step size of every 4 pixels, which is determined by the ratio of pixel size of the XRF data (30.0 nm) and probe (7.5 nm). Thus, the pixel size of enhanced XRF data matches the pixel size of reconstructed probe. Image patches of size 64×64 were randomly cropped from the low-resolution images, and their corresponding 256×256 patches were extracted from the high-resolution images. In total, 6280 patches were utilized for training. Data augmentation techniques were applied to these training images. They were normalized between 0 and 1, subjected to random rotations of 90°, 180°, and 270°, and could be flipped both horizontally and vertically. A total of 100 natural images from the DIV validation dataset were used for validation, all of which were normalized between 0 and 1. The prediction results can be seen in Fig. 1(b). We also demonstrate the improved importance of the dual-branch structure for feature extraction by conducting an ablation study (see Table S1 in Supplement 1).

C. Fine-Tuning the Machine-Learning Model with Experimental Data from Synchrotron X-Ray Microscopy

In the realm of image enhancement, models trained exclusively on natural images often encounter challenges when applied to real experimental data, especially with noise. Such models, when used to enhance these noisy experimental images, tend to introduce streak-like or periodic pattern artifacts such as the ones indicated by dashed lines shown in Fig. 4. This phenomenon can be attributed to the fact that natural images typically lack the specific types of noise present in experimental data [33]. The artifacts may come from different sources: (1) The noise condition we added in the simulation process could be different from the experimental condition. (2) The artifacts in the marked region in Fig. 4 show periodic line and dot structures, which share certain similarities with the ptychography reconstruction artifacts with periodic scan patterns. The intrinsic connection relies on periodic grids that were used to simulate the low-resolution images in the training process. (3) Overfitting in the de-convolution process may also introduce such artifacts. In this work, we show that fine-tuning the network with a small amount of experimental datasets can effectively mitigate those defects. Experimental data encompassed all kinds of noise and can enable the network to concurrently suppress all types of noise during the enhancement process, ensuring a comprehensive and optimized noise reduction strategy. As mentioned in experimental setup, two sets of datasets were measured, one with low resolution at 120 nm and high resolution at 30 nm, the other at 200 nm and 50 nm. Both have the same scale factor of four. The primary objective was to enhance the data at 30 nm resolution further by 4 times to match the quality of the ptychography reconstruction. Thus the dataset measured at 200 nm and 50 nm can be served to fine-tune the network. However, the features in the 50 nm XRF images, though we labeled them as "high-resolution," still seem to be blurry compared with the high-resolution natural images. As a result, it is inevitable to introduce a slightly blurry effect during the fine-tuning stage. It is important to strike a balance between the training from simulated natural images and fine-tuning so that the model can retain the excellent feature extraction and deblurring capabilities it acquired from training on natural images and suppress the noise. To achieve that goal, strategic adjustments were made during the fine-tuning stage: the learning rate was set at a tenth of that used for the natural image training, and the epoch count was limited to 50, a significant reduction from the 250 epochs designated for natural image training. Notably, the number of image patches employed for this fine-tuning process was 1024, deducted from 37 XRF images, which is considerably fewer than those used in the natural image training phase.

The network is designed to enhance its feature extraction capabilities, enabling it to effectively recover the fine features blurred in the scanning process. Furthermore, the fine-tuning stage, which

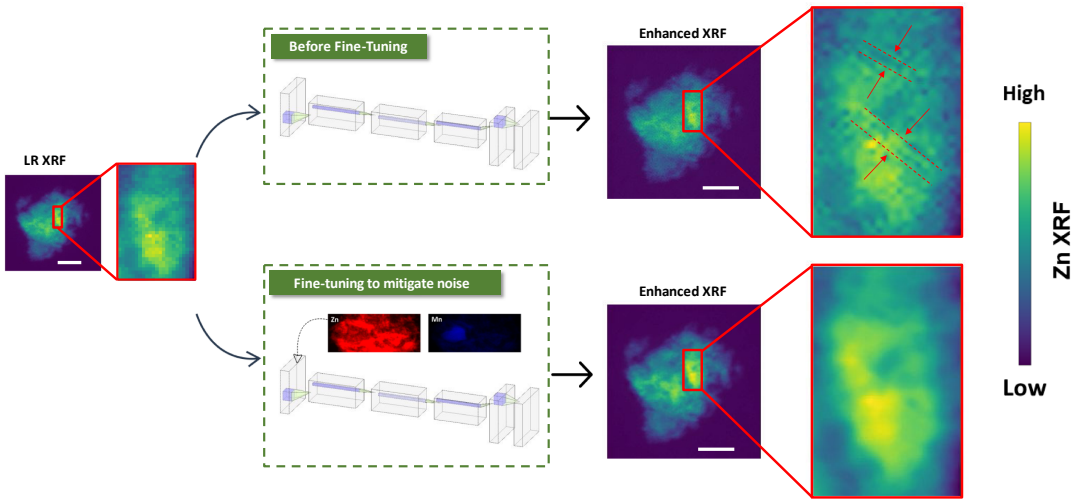

Fig. 4. Zn and Mn mapping at resolutions of 50 nm and 200 nm is used to fine-tune the network with a reduced learning rate, and fewer epochs, enabling the network to enhance the XRF projections with noise, effectively eliminating streak-like artifacts. Some of such artifacts are indicated by the arrows as examples. Zn XRF maps are shown as examples. All the scale bars are 1 μm.

helps to suppress the noise, bridges the domain gap between simulated data and experimental data, ensuring that the model performs robustly on specified experimental datasets.

4. RESULTS AND DISCUSSION

A. Enhancement on Lateral and Angular Sampling of XRF Data

Rechargeable aqueous batteries, especially Zn/MnO_2 batteries, have gained significant attention for large-scale energy storage due to their cost-effectiveness and inherent safety. Historically, Zn/MnO_2 batteries dominated the primary battery market, underscoring their affordability and safety [27]. In mild aqueous electrolytes, these batteries demonstrate promising results to be used as rechargeable batteries [28,34–36]. To mitigate these challenges and enhance cycling stability, understanding the nano-scale morphological and structural changes within the electrodes during the dissolution–deposition processes is essential. Assessing the impact of these changes on battery performance is vital for developing strategies to optimize battery longevity and efficiency.

For our study, a charged Zn/MnO_2 battery cathode in a mild aqueous electrolyte, after 450 h of cycling at a C-rate of 0.1C, was measured at the HXN beamline. Both XRF signals and far-field diffraction patterns were collected. The probe profile, reconstructed from ptychography phase retrieval, informed the network's learning to separate the blurring effect. Each XRF 2D image at 30 nm was enhanced. The raw XRF images had a pixel size of 30 nm and size $91 \times 140 \times 140$, with 91 representing the number of evenly distributed projections from 90° to −90°, and 140×140 indicating the projection's width and height. Based on the Crowther criterion, an increase in resolution necessitates a corresponding increase in the number of projections for valid tomographic reconstruction. Consequently, tomographic reconstructions from the 30 nm XRF projections were reprojected to 270 projections, evenly distributed from 90° to −90°. These reprojected projections, with a pixel size of 30 nm, were then enhanced by the proposed model with a scale factor of 4, resulting in enhanced XRF data with a pixel size of 7.5 nm, which matched the pixel size of the probe, and dimensions of $270 \times 560 \times 560$.

The tomographic reconstructions for both the raw and enhanced XRF were executed using the ordered-subset penalized maximum likelihood algorithm [37]. Figure 5 displays the comparison before and after enhancement for both 3D visualization and 2D central slices. Enhanced XRF reveals clearer particle borders, and even nano-scale cracks and pores are discernible. However, these minute features tend to blur during fly-scan. Fly-scan essentially averages the surrounding voxels, blurring boundaries. Trained on high- and low-resolution image pairs from natural image datasets and fine-tuned for experimental noise, the model effectively separates the blurring effect from each fly-scan spot. Using the Fourier shell analysis [38], the resolution of both raw and enhanced XRFs was determined. The intersection of the FSC with the one-bit threshold, as depicted in Figs. 5(d) and 5(e), provides the quantified resolution. The resolution of the raw XRF improved from 60.8 nm to 14.8 nm upon enhancement. In Fig. S3 in Supplement 1, we showed the relationship between number of reprojections and the calculated resolution by FSC. By re-projection and super resolution, 14.6 nm was the limit of this method to enhance the tomography reconstruction, and 270 projections met the minimum requirement to achieve that resolution. Such enhancement could enable quantitative analysis on finer structures, like cracks and pores, and offer insights into the formation of irreversible Zn/Mn complex side products during cycling. This understanding propels the Zn/MnO_2 system closer to practical application, highlighting its potential as a reversible, stable, and safe battery chemistry.

In Fig. 6, a comparison between the raw XRF, enhanced XRF, and phase-contrast image from ptychography reconstruction is shown to better validate the method. The enhanced XRF tomography reconstruction shows a more detailed structure, compared to the raw projections, which only offer pixelated features. This new, detailed structure in the enhanced projection correlates with the structure observed in the phase-contrast image obtained from ptychography reconstruction. A closer look reveals a porous structure within the sample, as indicated by the red box in Fig. 6, which remains unseen in the raw XRF due to blurry from fly-scan. However, this porous structure becomes visible in the enhanced XRF after decoupling the blurry effect from the scanning probe

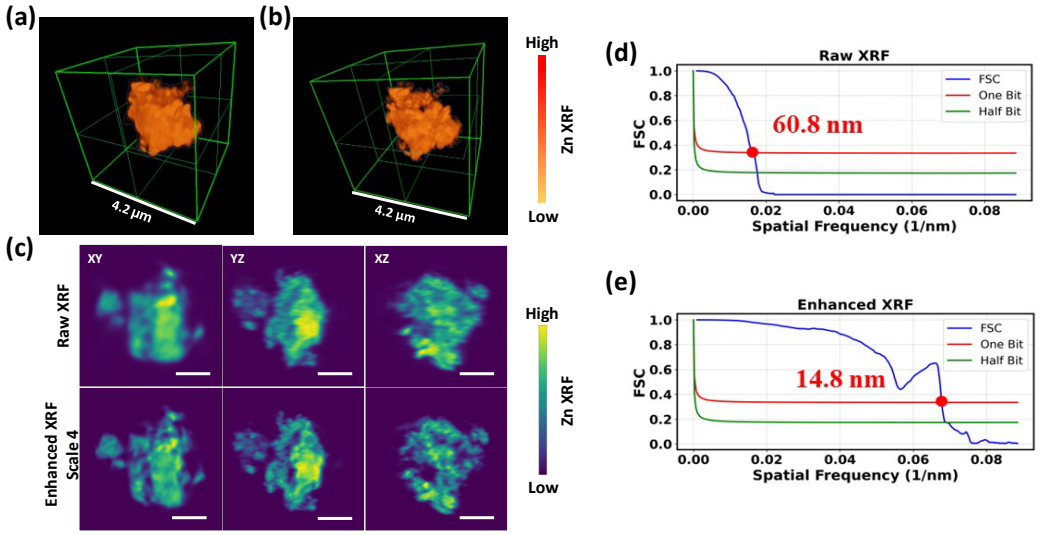

Fig. 5. Comparison of tomography reconstructions from raw (91 projections) and enhanced XRF images (270 projections). This includes central cross-sections in three directions and the application of the Fourier shell correlation (FSC) for resolution quantification. (a) and (b) display 3D visualizations of the raw and enhanced tomography reconstructions, respectively. (c) illustrates the cross-sections derived from both raw and enhanced XRF, highlighting the mitigation of blurring effects due to sample movement and the dynamic source, resulting in clearer representation of cracks and pores. (d) and (e) show the FSC analysis, with resolutions determined at 60.9 nm for raw and 14.8 nm for enhanced XRF.

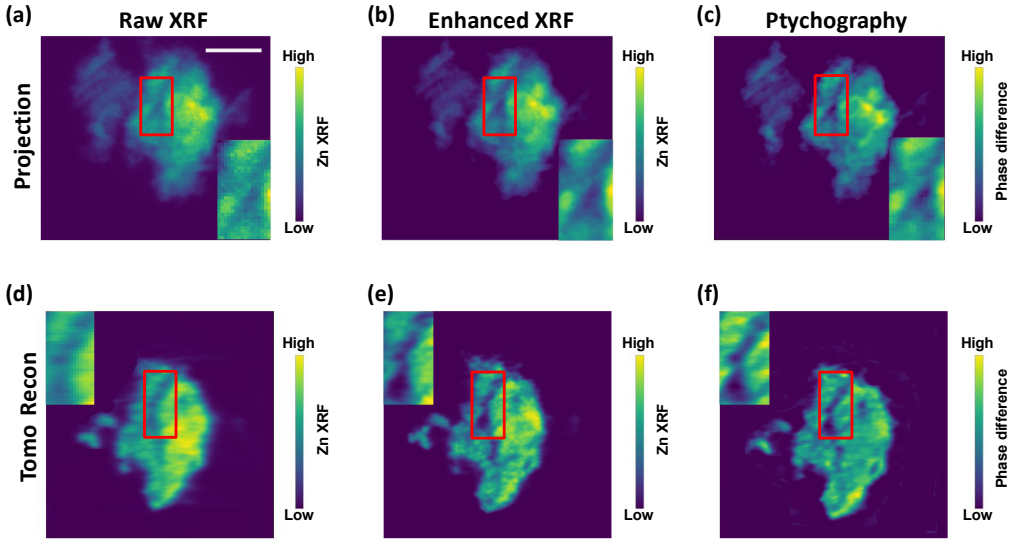

Fig. 6. Comparison between the enhanced XRF data and the phase-contrast image from ptychography reconstruction. (a) to (c) show the projections of raw XRF (fitted at Zn emission line), enhanced XRF (scale factor of 4), and phase-contrast image from ptychography reconstruction The zoom-in views of the labeled red box are at the bottom right corner of each image. (d) to (f) are one of the slices from tomography reconstruction with raw XRF data, enhanced XRF data, and ptychography data. The pixel size of those images is 30 nm (left column), 7.5 nm (middle column), and 7.5 nm (right column). The scale bar is 1 μm.

and is consistent with the ptychography reconstruction, while comparing the XRF fitted at Zn with the phase-image from ptychography reconstruction might not be entirely fair, as ptychography can reveal more morphological information beyond the distribution of Zn. Here, the ptychography serves as a validation method for the pore structure within the marked region, observable in the enhanced XRF but absent in the raw XRF data. It is anticipated that the phase-contrast ptychography and the enhanced XRF can be similar, though they may not be identical. However, the compound in our sample is $ZnMn_2O_4$ [27], thus Mn and Zn are co-localized within the sample. Therefore, the distribution of any given element, such as Zn, could indicate the primary structural information in comparison to the phase-contrast image.

B. Enhancement on Tomography Reconstructions with Limited Projections

Data acquisition at the HXN beamline, especially with a fine step size, is notably time-consuming. For instance, obtaining 91 projections at a 140 × 140 pixel size field of view at 30 nm step size requires around 16 h. This slow rate of data acquisition makes it impractical to conduct a tomography measurement with sufficient projections determined by the Crowther criterion. Considering the inherent dependency of projections at adjacent angles and the well-trained super resolution model, it is feasible to obtain a valid high-resolution tomography reconstruction with a reduced number of projections. To mimic a real-world situation where

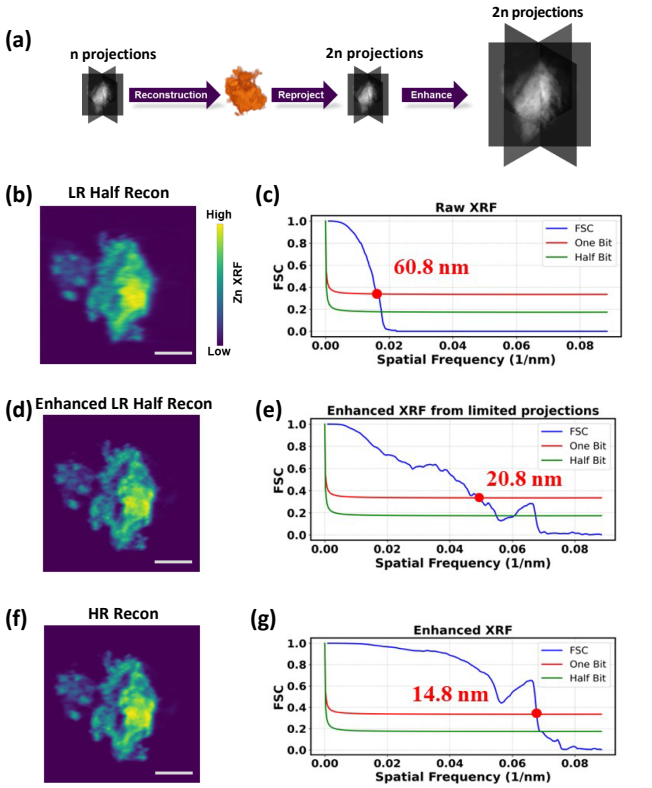

Fig. 7. Results of limited angle tomography reconstruction. (a) Illustrates the reprojection process for enhancement in the angular dimension. Half of the low-resolution Zn XRF projections (45 in total) are used for tomography reconstruction. This reconstructed volume is subsequently reprojected across 91 angles, uniformly spanning from −90° to 90° to obtain projections at the same angle as the experiment. These reprojected Zn XRFs are enhanced by the model to achieve high-resolution Zn XRFs, which were then used for tomography reconstruction for the 3D Zn distribution. (b) and (d) show a central slice from the raw and enhanced tomography reconstructions in the YZ plane, respectively. (c) and (e) present the FSC analysis, quantifying the resolution of the raw and enhanced tomography reconstructions. Resolutions are 60.8 nm for the raw and 20.8 nm for the enhanced tomography reconstruction. For comparison, (f) and (g) show the central slice from the tomography reconstruction with the original 91 projections and the FSC analysis, which indicates a resolution of 14.8 nm. All the scale bars are 1 μm.

limited projections were available, only half of the projections evenly distributed from 90° to −90° acquired at 30 nm were used for tomography reconstruction. Following this, the tomographic reconstruction was reprojected into 91 projections to have a fair comparison of the real experimental XRF data with the same number of projections. The schematic of the reprojection process is shown in Fig. 7(a). The proposed trained model was then employed to enhance these projections, after which tomography reconstruction was performed using the enhanced 91 projections. The cross sections in Figs. 7(b) and 7(d) indicate that this method can still yield a valid reconstruction of the sample, thereby accelerating the experiment. The resolution for raw and enhanced XRF using half of the projections is 60.8 nm and 20.8 nm by applying the enhancement on the 91 measured projections. It is noteworthy that the resolution of 20.8 nm represents a nearly 3-fold improvement over the 60.8 nm obtained from the raw XRF data. This underscores the validity of the assumption regarding the similarity between adjacent projections and attests to the model's capability to enhance the resolution of low-quality XRF data. Moreover, the 20.8 nm resolution closely approximates the 14.8 nm achieved by inserting more projections for tomographic reconstruction. Yet, this enhanced resolution can be achieved in half the experimental duration. This also suggests that incorporating additional projections could further refine the resolution according to the Crowther criterion. Additionally, this refined resolution effectively unveils nuanced nano-scale structures, such as cracks and pores, facilitating a more precise morphological and chemical analysis pertinent to the study of reaction mechanisms in energy storage research.

5. CONCLUSION

In this research, a dual-branch network enabling feature extraction at both space and frequency domains was employed to enhance the resolution of XRF images. This network was trained using fly-scan simulations on natural images to decouple the blurring from fly-scan during data acquisition. It was subsequently fine-tuned with a reduced learning rate and limited epochs using real experimental data to suppress noise. The enhancement was done both on lateral sampling in each projection by the proposed model and on angular sampling by inserting new projections. The resolution of the enhanced XRF tomography reconstruction is enhanced by 4 times as indicated by FSC, compared to the raw XRF data. More distinct and fine features like cracks and pores are visible in the enhanced XRF data, which helps to elucidate the underlying reaction mechanisms inherent to energy storage materials. Furthermore, by fully capitalizing on the inherent relationship between adjacent projections and the well-trained model, valid enhanced XRF tomography reconstructions could be obtained with only half of the low-resolution XRF projections. Overall, the proposed dual-network super-resolution imaging method based on deep-learning can enable high-resolution tomography reconstruction by enhancement in the lateral sampling of each 2D images and also in the angular sampling by inserting new 2D images. Such an approach not only conserves experimental time but also minimizes potential beam damage to the sample.

Funding. U.S. Department of Energy, Office of Science, Basic Energy Sciences (DE-SC0012673, DE-SC0012704).

Acknowledgment. We thank the Founding Director, Prof. Esther Takeuchi, and the Center Director, Prof. Amy Marschilok, of the Center for Mesoscale Transport Properties EFRC for constructive input regarding the sample of interest. We thank the HXN beamline team members for the instrumentation development. Part of the student effort from X. Zheng was supported by the Imaging and Microscopy Program at the National Synchrotron Light Source II of Brookhaven National Laboratory. This motivation for the selected sample and the associated sample preparation and electrochemistry were supported as part of the Center for Mesoscale Transport Properties (m2m#S), an Energy Frontier Research Center (EFRC) supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under award No. DE-SC0012673. This research used resources and Hard X-ray Nanoprobe Beamline (HXN, 3-ID) of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. M. D. De Jonge and S. Vogt, "Hard X-ray fluorescence tomography—an emerging tool for structural visualization," Curr. Opin. Struct. Biol. **20**, 606–614 (2010).
2. C. M. Krauland, L. C. Jarrott, R. P. Drake, *et al.*, "An evaluation of high energy bremsstrahlung background in point-projection x-ray radiography experiments," Rev. Sci. Instrum. **83**, 10E528 (2012).
3. H. Yan, N. Bouet, J. Zhou, *et al.*, "Multimodal hard x-ray imaging with resolution approaching 10 nm for studies in material science," Nano Futures **2**, 011001 (2018).
4. E. Nazaretski, H. Yan, K. Lauer, *et al.*, "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 (2017).
5. X. Huang, K. Lauer, J. N. Clark, *et al.*, "Fly-scan ptychography," Sci. Rep. **5**, 9074 (2015).
6. H. Wang, R. C. Atwood, M. J. Pankhurst, *et al.*, "High-energy, high-resolution, fly-scan X-ray phase tomography," Sci. Rep. **9**, 8913 (2019).
7. Y. Jiang, J. Deng, Y. Yao, *et al.*, "Achieving high spatial resolution in a large field-of-view using lensless x-ray imaging," Appl. Phys. Lett. **119**, 124101 (2021).
8. P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," Nature **494**, 68–71 (2013).
9. J. J. Deng, D. J. Vine, S. Chen, *et al.*, "X-ray ptychographic and fluorescence microscopy of frozen-hydrated cells using continuous scanning," Sci. Rep. **7**, 445 (2017).
10. D. J. Vine, D. Pelliccia, C. Holzner, *et al.*, "Simultaneous X-ray fluorescence and ptychographic microscopy of *Cyclotella meneghiniana*," Opt. Express **20**, 18287–18296 (2012).
11. J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *14th European Conference on Computer Vision (ECCV)* (2016).
12. C. Dong, C. C. Loy, K. He, *et al.*, "Image super-resolution using deep convolutional networks," arXiv, arXiv:1501.00092 (2015).
13. S. Gu, W. Zuo, Q. Xie, *et al.*, "Convolutional sparse coding for image super-resolution," in *IEEE International Conference on Computer Vision (ICCV)* (2015).
14. A. M. Maiden, M. J. Humphry, F. Zhang, *et al.*, "Superresolution imaging via ptychography," J. Opt. Soc. Am. A **28**, 604–612 (2011).
15. B. Lim, S. Son, H. Kim, *et al.*, "Enhanced deep residual networks for single image super-resolution," arXiv, arXiv:1707.02921 (2017).
16. S. Gao and X. Zhuang, "Multi-scale deep neural networks for real image super-resolution," arXiv, arXiv:1904.10698 (2019).
17. L. Wu, S. Bak, Y. Shin, *et al.*, "Resolution-enhanced X-ray fluorescence microscopy via deep residual networks," npj Comput. Mater. **9**, 43 (2023).
18. Y. Zhang, K. Li, K. Li, *et al.*, "Image super-resolution using very deep residual channel attention networks," in *European Conference on Computer Vision (ECCV)*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, eds. (Springer, 2018), pp. 294–310.
19. J. Li, F. Fang, K. Mei, *et al.*, "Multi-scale residual network for image super-resolution," in *European Conference on Computer Vision (ECCV)*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, eds. (Springer, 2018), pp. 527–542.
20. E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: dataset and study," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, 2017), pp. 1122–1131.
21. J. Sherman, "The theoretical derivation of fluorescent X-ray intensities from mixtures," Spectrochim. Acta **7**, 283–306 (1955).
22. K. Li, Z. Wu, K.-C. Peng, *et al.*, "Guided attention inference network," IEEE Trans. Pattern Anal. Mach. Intell. **42**, 2996–3010 (2020).
23. F. Wang, M. Jiang, C. Qian, *et al.*, "Residual attention network for image classification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017), pp. 6450–6458.
24. Y. Tai, J. Yang, X. Liu, *et al.*, "MemNet: a persistent memory network for image restoration," in *IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2017), pp. 4549–4557.
25. Y. Tai, J. Yang, and X. Liu, "Image super-resolution via deep recursive residual network," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017), pp. 2790–2798.
26. W. Shi, J. Caballero, F. Huszár, *et al.*, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," arXiv, arXiv:1609.05158 (2016).
27. V. R. Kankanallu, X. Zheng, D. Leschev, *et al.*, "Elucidating a dissolution–deposition reaction mechanism by multimodal synchrotron X-ray characterization in aqueous Zn/MnO_2 batteries," Energy Environ. Sci. **16**, 2464–2482 (2023).
28. H. Pan, Y. Shao, P. Yan, *et al.*, "Reversible aqueous zinc/manganese oxide energy storage from conversion reactions," Nat. Energy **1**, 16039 (2016).
29. A. F. Agarap, "Deep learning using Rectified Linear Units (ReLU)," arXiv, arXiv:1803.08375 (2019).
30. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv, arXiv:1412.6980 (2017).
31. H. Zhao, O. Gallo, I. Frosio, *et al.*, "Loss functions for neural networks for image processing," arXiv, arXiv:1511.08861 (2018).
32. A. Paszke, S. Gross, S. Chintala, *et al.*, "Automatic differentiation in PyTorch," in *31st Conference on Neural Information Processing Systems (NIPS 2017)* (Curran Associates Inc., 2017).
33. B. Beckhoff, *Handbook of Practical X-Ray Fluorescence Analysis* (Springer, 2006).
34. D. Wu, L. M. Housel, S. T. King, *et al.*, "Simultaneous elucidation of solid and solution manganese environments via multiphase operando extended X-ray absorption fine structure spectroscopy in aqueous Zn/MnO_2 Batteries," J. Am. Chem. Soc. **144**, 23405–23420 (2022).
35. D. Wu, L. M. Housel, S. J. Kim, *et al.*, "Quantitative temporally and spatially resolved X-ray fluorescence microprobe characterization of the manganese dissolution-deposition mechanism in aqueous Zn/α-MnO_2 batteries," Energy Environ. Sci. **13**, 4322–4333 (2020).
36. X. Zuo, K. Chang, J. Zhao, *et al.*, "Bubble-template-assisted synthesis of hollow fullerene-like MoS_2 nanocages as a lithium ion battery anode material," J. Mater. Chem. A **4**, 51–58 (2016).
37. D. a. Gürsoy, F. De Carlo, X. Xiao, *et al.*, "TomoPy: a framework for the analysis of synchrotron tomographic data," J. Synchrotron Radiat. **21**, 1188–1193 (2014).
38. M. van Heel and M. Schatz, "Fourier shell correlation threshold criteria," J. Struct. Biol. **151**, 250–262 (2005).