Fully automated deep learning based resolution recovery

Matthew Andrew*[a], Andriy Andreyev[a], Faguo Yang[a], Lars Omlor[a], Masako Terada[a], Allen Gu[a], Robin White[a]
[a]Carl Zeiss X-ray Microscopy, Dublin, CA USA, 94558

ABSTRACT

A novel automated workflow for the recovery of image resolution using deep convolutional neural networks (CNNs) trained using spatially registered multiscale data is presented. Spatial priors, coupled with high order voxel-based image registration, are used to correct for uncertainties in image magnification and position. A network is then trained to remove the effects of point spread from the low-resolution data, improving resolution while reducing image noise & artefact levels. While benchmarking on real materials, including biological, materials science and electronics samples, we find that resolution recovery improves quantitative and qualitative measurements, even if certain image details cannot be easily identified from the original low-resolution data.
Keywords: Computed Tomography, Image Reconstruction, Deep Learning, self-supervised learning, AutoML.

1. INTRODUCTION

One of the principal challenges in tomographic imaging is image sharpness. Effects such as system instability, detector blurring, and a finite source spot size can induce imaging artefacts. This is compounded by partial volume effects, where material interfaces lie within the volume described by a voxel, leading to a reconstructed intensity to spill over material phases. These effects can be modelled as a Point Spread Function (PSF) convolved with a structure function. The removal of a PSF from X-ray CT data is extremely challenging, however, as the definition of the PSF is strongly dependent on system / analytical conditions, requires specific measurement, and typically cannot account for spatially variable effects (e.g. due to beam hardening, PSF anisotropy or scatter). The removal via PSF deconvolution can also greatly increase image noise and be computationally expensive.

In this paper we present a novel workflow for the recovery of image resolution using a deep convolutional neural network (CNN) trained using spatially registered multiscale data. Spatial priors, coupled with high order voxel-based image registration is used to correct for residual uncertainties in image magnification and position. A network is then trained to remove the effects of point spread from the low-resolution data, recovering blurred features and aiding in image segmentation and analysis. This network effectively learns the coupled effects of a finite PSF and partial volume artifacts present in the low resolution data by correlating it to high resolution data. As sources of spurious signal (noise and sampling artefacts) are uncorrelated between the low- and high-resolution data, the network also has the effect of removing these noise and artefacts from the recovered image, in contrast to traditional techniques for PSF removal which typically amplify noise.

This approach is qualitatively and quantitatively benchmarked across a range of samples, including plant science, energy materials and electronic samples. In all cases small structures are effectively recovered. Quantitative measurements also find that structural properties (such as porosity, pore size, pore number, and pore volume), critical for material characterization, are accurately predicted after the use of the resolution recovery network even if they were not accurately predicted from the original low-resolution data.

2. METHODS

System image blur arises from a range of sources, including source spot size, sample stability limits, scintillator light spread, and optical / diffraction resolution limitations. Traditional approaches for the reduction of this image blur center around the concept of spatial PSF deconvolution whereby the whole system PSF (usually described by its Fourier transform the Modulation Transfer Function (MTF)) is deconvolved from the detected image. The practical application of this technique, however, faces many challenges. MTF determination requires significant approximations / simplifications based on analytical geometry definitions and / or fitted calibration phantoms which may not be appropriate for use in real samples. The deconvolution process is often slow and computationally expensive, it typically increases image noise (by up-weighting high frequency components). Finally, the MTF is usually assumed to be

7th International Conference on Image Formation in X-Ray Computed Tomography,
edited by Joseph Webster Stayman, Proc. of SPIE Vol. 12304, 123041Q
 · 0277-786X · doi: 10.1117/12.2647272

spatially invariant, making it non-adaptive to anisotropic effects such as scatter or beam hardening. In this paper we prepose a novel approach to image resolution recovery whereby this deconvolution step is replaced a convolutional neural network trained specifically on the low-to-high feature map specific for the structures being imaged.

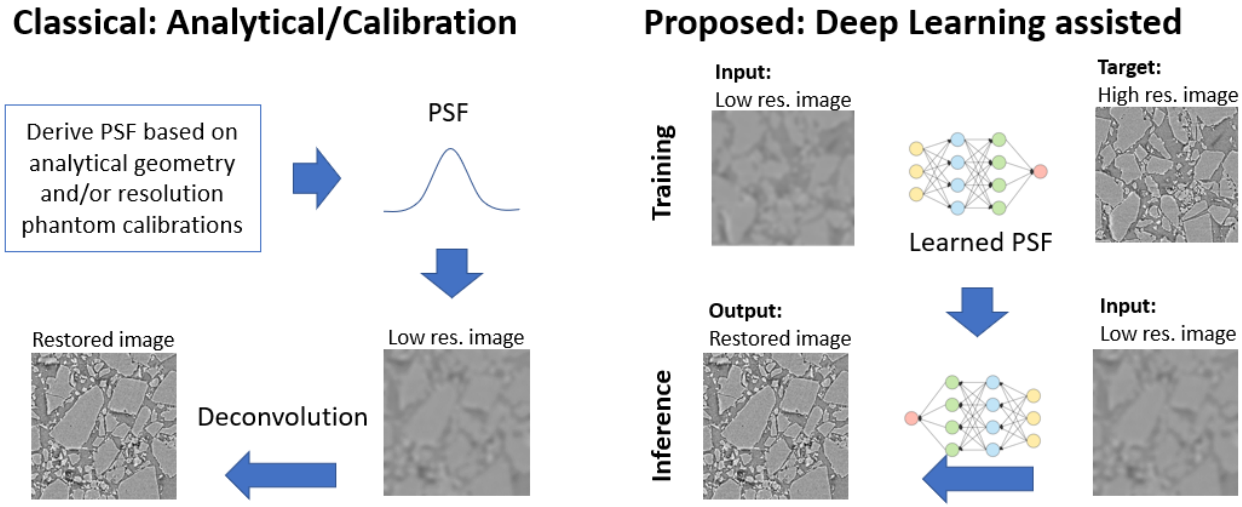

Figure 1: Contrasting traditional (analytical) PSF removal to Deep Learning based resolution recovery.

The fundamental advantage of X-ray imaging, relative to other microanalytical techniques, is that it is non-invasive. This allows for the imaging of high-resolution regions-of-interests (ROIs) within large objects, and the relative stage coordinates used for this workflow allow for strong spatial priors for multiscale image registration techniques [1], [2]. Such multiscale imaging is often challenging in laboratory micro-CT systems as the requirement for geometric magnification from a point source of X-rays limits the magnification attainable on large objects within a limited system geometry. The use of objective-coupled scintillators allows for extremely small pixels on the imaging (scintillator) plane, and correspondingly high resolutions even at limited geometric magnifications. These "Scout-and-Zoom" workflows have been broadly used for the targeting of high-resolution regions-of-interest based on macroscopic context, however the ability to use multiscale data quantitatively, even when registered perfectly to sub-pixel precision, has, to date, been limited.

This paper presents a novel workflow for the recovery of small-scale structures from larger scale data. The use of multiscale data allows resolution recovery to be automatically tailored to a sample class and acquisition settings and removes any challenging "calibration" steps. First a multiscale dataset is acquired, consisting of a low resolution, Large Field-of-View (LFOV) data and a high-resolution smaller ROI data entirely contained in the volume of the LFOV image. Partial pixel dithering is used on both acquisitions, both to remove image artefacts and to encode sub-resolution information into the projection dataset. The stage positions of the high-resolution ROI are used to calculate the volume of interest within the LFOV volume, and a sub-region of the LFOV image is directly reconstructed to the smaller voxel size of the high-resolution data using sub-pixel FDK reconstruction [3]. While, for a well aligned system, the ROI and high-resolution reconstructions will be well aligned, residual offsets (due to imperfect alignment in the optical train and uncertainties in stage position) lead to small offsets between the different resolutions and data positions. AI based image-to-image regression is extremely sensitive to image offsets (e.g. [4]), and so the images must first be registered to sub-pixel precision. This was performed using a multiscale registration process using a normalized Matte's mutual information image metric [5], a LBFGS optimizer [6], and a translation transform in the first step, followed by a full 9 degree of freedom transform in the second. The transformed LFOV subpixel image was then interpolated onto the voxel grid of the high-resolution dataset, creating a well-formed image regression match.

Resolution recovery networks were then trained using a modification of well-established image-to-image regression techniques [7], [8], with loss functions, network structures and data augmentation tailored to the 3D tomographic imaging problem, modified to ensure rapid convergence and high performance even with early stopping [9]. Network structures were adapted from the U-Net architecture [10]. Both noise and most artefacts were uncorrelated between the input and target datasets (because of measurement independence and differing system geometry respectively), while the

LFOV to High Resolution structures (and the features that predict those structures) are strongly correlated. This causes networks to learn both this feature map and remove noise and artefacts simultaneously, with noise and artefact removal occurring due to a process conceptually analogous to Noise-to-Noise denoising [7]. Data was dynamically augmented during training for variations in intensity and orientation during training, improving model robustness. Once model training is complete, inference can be performed across the entire LFOV image, as reconstructed onto the high-resolution image grid, or specific regions of interest within the LFOV image.

Image training is performed in a Python-based TensorFlow 2.8 environment, after which network weights were transferred into an .onnx format and inference performed using NVIDIA TensorRT 8.0 runtime. An already trained model can also be applied to any other structurally similar sample if the acquisition conditions still match the conditions for which the model has been trained for. This greatly expands effective high-resolution field of view and improves throughput required to image multiple samples at high-resolution.

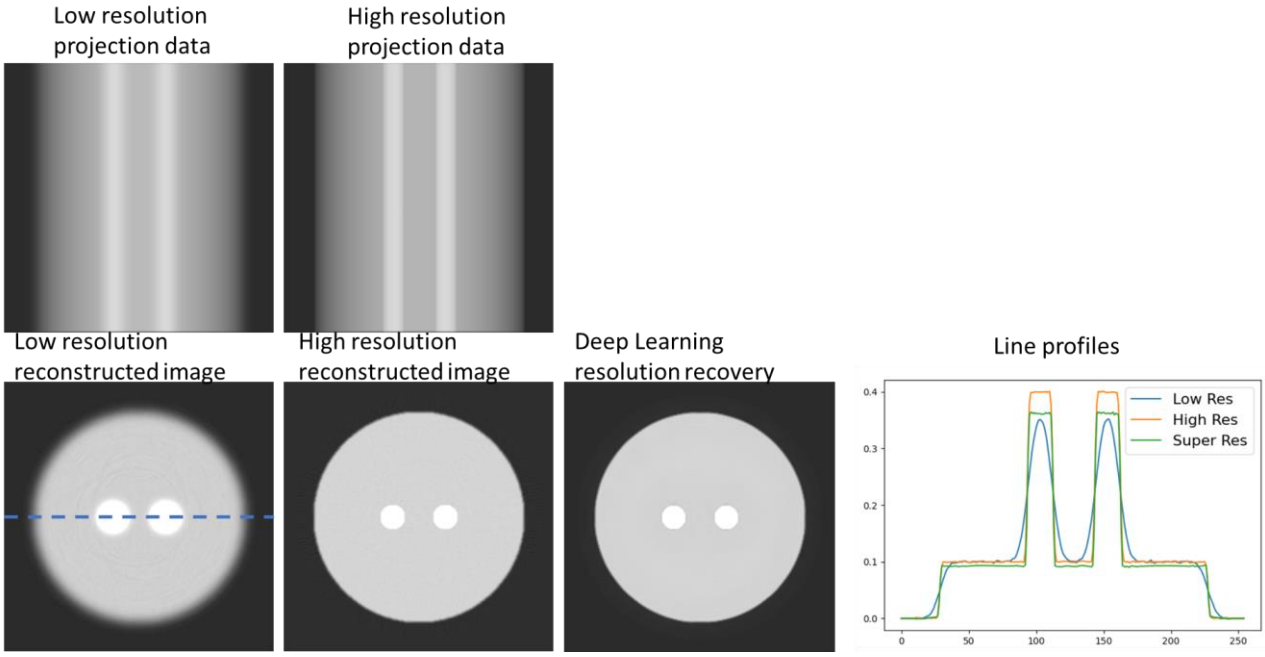

Figure 2. Proof of principle (sanity check) of the proposed Deep Learning assisted resolution recovery. Here, the low resolution image is used as an input, and high resolution image serves as target, neural network is able to near perfectly learn the mapping (in other words, point spread function) between the low resolution and high resolution data.

3. RESULTS AND DISCUSSION

This technique was benchmarked on a range of samples, with both quantitative and qualitative performance metrics analyzed from inside and outside the field of view used for training. The acquisition settings for all samples are shown in table 1.

First, the multiscale imaging workflow described above was applied to a 1-2 mm tomato seed. This sample was chosen to show the performance of the workflow on typical life science samples, with structures typical to life science systems, as well as provide a feature rich system to evaluate the performance of the resolution recovery workflow as the imaging approaches the limiting resolution of the system. The high-resolution data showed distinct cell boundaries and clear sub-cellular features, but at a much more constrained FOV. After resolution recovery the subcellular features are much clearer and cell boundaries are distinct, even in portions of the sample outside of the training region.

Sample	Tomato seed	Cathode particle	A12 Chip
System	Versa XRM	Ultra XRM	Context micro-CT
Tube voltage, kVp (low res / high res)	80 / 40	N/A (monochromatic 5.4keV)	160 / 120
Detector (low res / high res)	4X / 4X	LFOV objective / HRES objective	FPX / FPX
Projection array (low res / high res)	2048^2 / 2048^2	1024^2 / 512^2	3072^2 / 1536^2
Number of Projections (low res / high res)	2401 / 4001	901 / 901	1601 / 2401
Voxel size, µm (low res / high res)	1.7 / 0.7	0.065 / 0.032	10.35 / 2.1
Volumetric FOV, mm^3 (low res / high res)	44.2 / 2.9	0.00027 / 0.0000041	20,000 / 170
Acquisition time, hrs (low res / high res)	2 / 14	12 / 40	1.5 / 2.5
Volume adjusted throughput improvement, X	100	215	200

Table 1. Acquisition settings for the sample images presented in this paper.

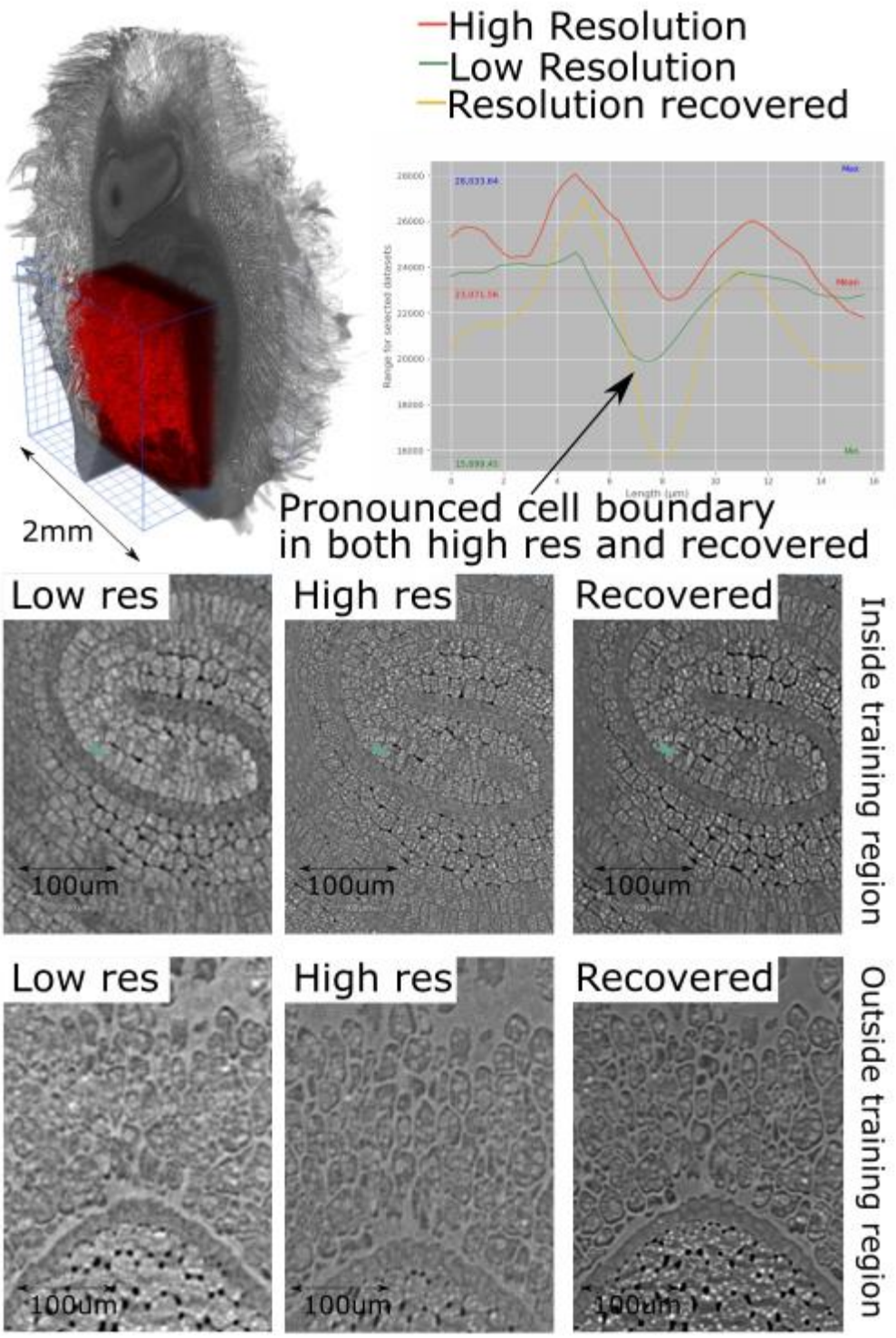

Figure 3: Resolution recovery on a tomato seed sample, showing significant boosts in image quality both inside and outside the training region.

The next example application is a set of battery cathode particles, where porosity, pore size and distribution is critical to understand battery performance and degradation. The sample was scanned at very high resolution using ZEISS Xradia Ultra XRM, which is a zone plate microscope with two objectives. First, a “Large Field of View” objective with an

effective voxel size of 65nm, a field of view of around 65µm, and a spatial resolution of around 100nm, and also a high resolution objective with a voxel size of 32nm, a spatial resolution of 50nm and a field of view of 16µm. Three scans were performed, one with the LFOV objective, of the entire particle set, taking 12 hours, and two individual particles were scanned at high resolution, taking 40 hours each (figure 4). The LFOV scans had a volumetric field of view equivalent to 64x the volumetric field of view on the high resolution scans, equivalent of a volume-normalized throughput boost of 215X. The same procedure for multiscale 3D volumetric registration followed by network training and inference was followed. Quantitative analysis of the cathode particle porosity was also performed, shown below in table 2.

Measure	LFOV	HRES	Recovered
Whole particle volume, μm^3	1137	1137	1137
Porosity	3.6%	2.4%	2.8%
Total pore count	1154	3055	2600
Pore Volume, um^3	40.7	28.4	33.0

Table 2. Analysis results from cathode particle dataset.

Whole (filled) particle volume measurement for the three techniques agreed very closely, due the relatively large size of the individual particle (approximately 10-15um). This also shows that the image segmentation routine used for each dataset introduced very little bias between the measurements. For each other measurement, the resolution recovered volume measured much closer to the high resolution dataset than the original large field of view image, while offering a greatly expanded field of view relative to the high resolution dataset.

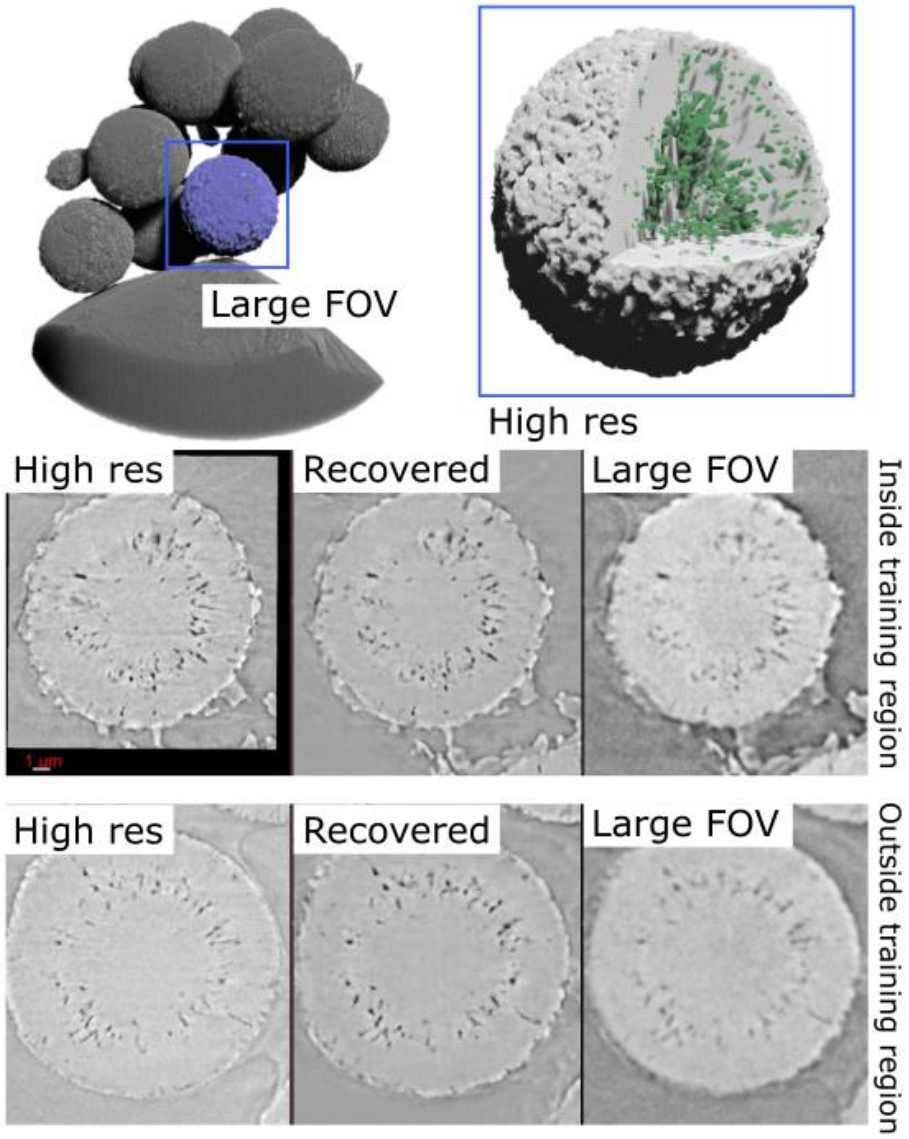

Figure 4: Resolution recovery on battery cathode particle dataset. High resolution porosity is more clearly visible in the resolution recovered dataset than in the original Large FOV datasets, even outside of the training region.

The final example is a commercially acquired A12 ARM system-on-a-chip (SOC), with the package targeted for an ROI (a sample typical to semiconductor failure analysis (FA) applications). In this application field of view is critical as there may be large arrays of solder bumps to be which need to be examined for defects, misalignments or cracks. The sample was thermally cycled 1,000 times to force fatigue and induce cracking according to the JEDEC standard. It was scanned at a large field of view which allowed for the entire package solder bump array to be seen, the high resolution scan of package solder bumps acquired to show specific crack distributions. The cracks / defects were easily visible on both the high resolution and resolution recovered datasets, but were obscured in the original large FOV dataset (figure 5).

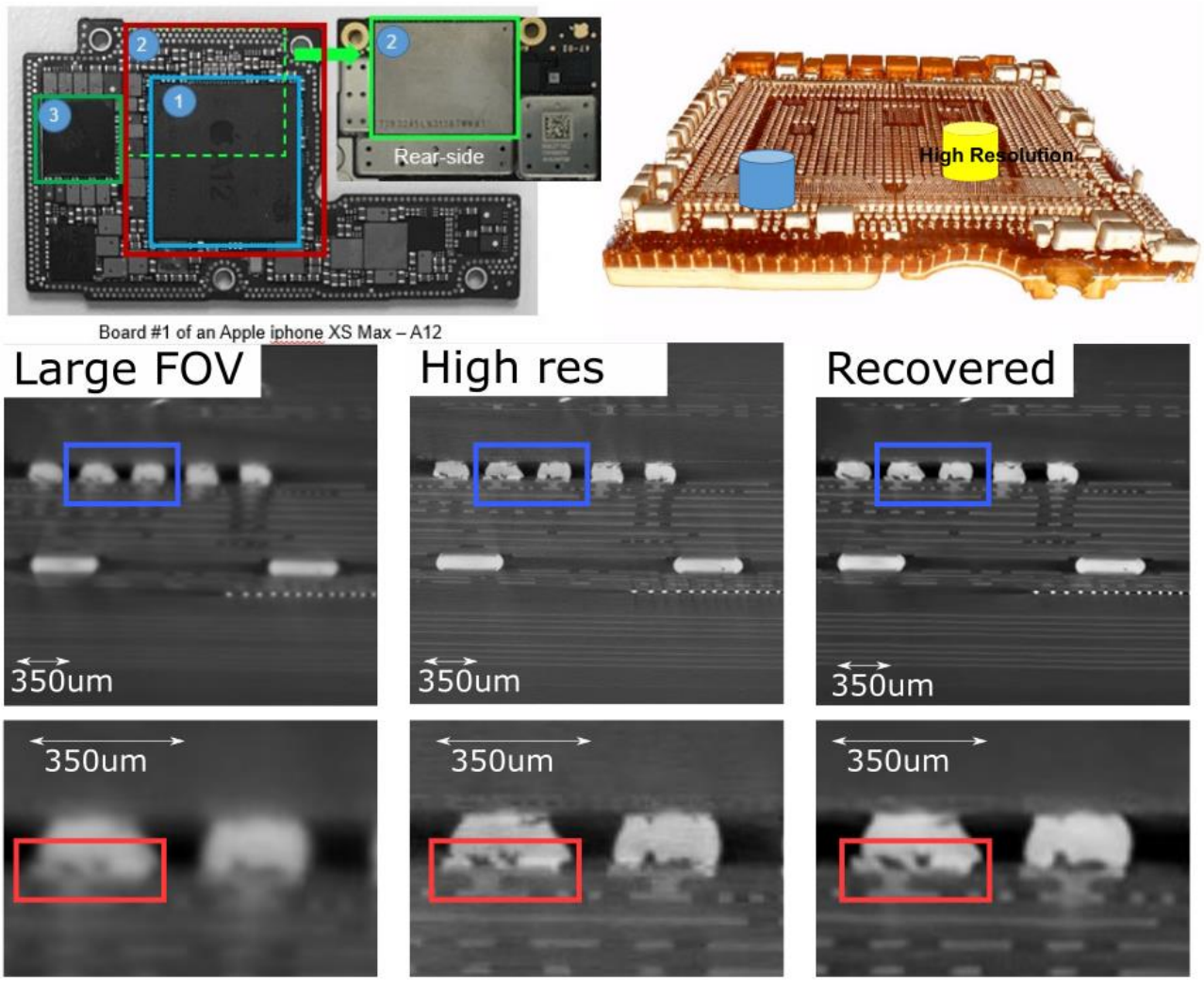

Figure 5: A12 chip scan, with the board position show with the first box (top left). A 3D rendering shows high resolution positions (top right). 2D slices of the large FOV dataset, high resolution and resolution recovered datasets are shown, as well as digital zooms of two solder bumps with a region of interest indicated by the blue box. The fracture region is shown with the red box. The fractures are evident in both high resolution and resolution recovered data, but not in the large FOV data.

4. CONCLUSIONS

A novel workflow for automated resolution recovery using a combination of multiscale imaging, automated registration and AI based PSF deconvolution is presented. This is benchmarked on a range of samples, including a tomato seed, cathode particle and system-on-a-chip. For each of these samples, image quality was dramatically improved without sacrificing field of view or throughput. In the case of the tomato seed, subcellular features became much clearer and cell boundaries become much more distinct with an effective throughput boost of around 100X. In the case of the cathode particle quantitative analyses showed much closer agreement to high resolution results after resolution recovery than before. In the case of the system on a chip fractures were apparent in the resolution recovered dataset, whereas they were not clear on the original LFOV dataset. The tradeoff between resolution and field of view is one of the fundamental limitations in microscopy and microanalysis. Imaging technologies such as AI based resolution recovery and PSF removal have the potential to break these fundamental limitations, allowing for much larger fields of view, throughput and novel applications for X-ray imaging.

REFERENCES

[1] J. Gelb, D. P. Finegan, D. J. L. Brett, and P. R. Shearing, "Multi-scale 3D investigations of a commercial 18650 Li-ion battery with correlative electron- and X-ray microscopy," *J. Power Sources*, vol. 357, 2017, doi: 10.1016/j.jpowsour.2017.04.102.

[2] M. G. Andrew, R. Chicha-Szot, S. Liden, P. Such, G. Lesniak, and A. Wiegmann, "Non-Invasive Multi-Scale Imaging and Modelling Using X-Ray Microscopy," *Microsc. Microanal.*, vol. 22, no. 108, 2016.

[3] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A*, vol. 1, no. 6, p. 612, 1984, doi: 10.1364/JOSAA.1.000612.

[4] R. Mechrez, I. Talmi, and L. Zelnik-Manor, "The contextual loss for image transformation with non-aligned data," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11218 LNCS. doi: 10.1007/978-3-030-01264-9_47.

[5] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellyn, and W. Eubank, "Nonrigid multimodality image registration," *Proc. SPIE Med. Imaging*, vol. 4322, 2001.

[6] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.*, vol. 45, no. 1–3, 1989, doi: 10.1007/BF01589116.

[7] J. Lehtinen *et al.*, "Noise2Noise: Learning image restoration without clean data," in *35th International Conference on Machine Learning, ICML 2018*, 2018, vol. 7.

[8] S. J. Jackson, Y. Niu, S. Manoorkar, P. Mostaghimi, and R. T. Armstrong, "Deep Learning of Multiresolution X-Ray Micro-Computed-Tomography Images for Multiscale Modeling," *Phys. Rev. Appl.*, vol. 17, no. 5, p. 54046, May 2022, doi: 10.1103/PhysRevApplied.17.054046.

[9] M. Andrew, L. Omlor, A. Andreyev, ravikumar Sanapala, and M. Samadi Khoshkhoo, "New technologies for x-ray microscopy: phase correction and fully automated deep learning based tomographic reconstruction," 2021. doi: 10.1117/12.2596592.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation [2015; First paper exploring U-Net architecture.]," 2015.