Machine learning for the reconstruction and analysis of synchrotron-radiation tomography data

Julian Moosmann[a], Jennifer Ahrens[b], Sarah Irvine[a], Tak Ming Wong[a], Christian Lucas[a,d], Felix Beckmann[a], Jörg U. Hammel[a], D.C. Florian Wieland[a], Berit Zeller-Plumhoff[a,c], and Philipp Heuser[b]

[a]Helmholtz-Zentrum Hereon, Max-Planck-Straße 1, 21502 Geesthacht, Germany
[b]Helmholtz Imaging, Deutsches Elektronen Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany
[c]University of Rostock, Albert-Einstein-Str. 2, 18059 Rostock, Germany
[d]Bruker Daltonics SPR, Falkenried 88, 20251 Hamburg, Germany

ABSTRACT

The Helmholtz-Zentrum Hereon is operating imaging beamlines for X-ray tomography (P05 IBL, P07 HEMS) for academic and industrial users at the synchrotron-radiation source PETRA III at DESY in Hamburg, Germany. The high flux density and coherence of synchrotron radiation enable high-resolution *in situ/operando/in vivo* tomography experiments and phase-contrast imaging techniques, respectively. Large amounts of 3D and 4D data are collected that are difficult to process and analyze. Recently, we have explored machine learning approaches for the reconstruction, processing and analysis of synchrotron-radiation tomography data. Here, we report on the application of supervised learning for multimodal data analysis to generate a virtual 3D histology, digital volume correlation of 4D *in situ* tomography data, and instance segmentation. Furthermore, we present findings related to unsupervised learning in the context of semantic segmentation.

Keywords: machine learning, X-ray imaging, tomography, synchrotron radiation, semantic segmentation, instance segmentation, multimodal imaging, histology, digital volume correlation

1. INTRODUCTION

The Helmholtz-Zentrum Hereon is operating several end stations for X-ray tomography at the P05 imaging beamline (IBL) and the P07 high energy material science beamline (HEMS) for academic and industrial users at the synchrotron-radiation source PETRA III at DESY in Hamburg, Germany. The high flux density and coherence of the synchrotron radiation at PETRA III enable high-resolution *in situ/operando/in vivo* tomography experiments and phase-contrast imaging techniques for weakly attenuating objects or to enhance intensity contrast, respectively. Large amounts of 3D and 4D data are routinely collected from a wide variety of samples with different spatio-temporal resolutions, noise levels, and artifacts, which is challenging to process and analyze. Recently, we have developed and applied various machine learning approaches for the processing, reconstruction, and analysis of micro- and nanotomography data. For the semantic segmentation of tomography data exhibiting a highly heterogeneous texture, supervised learning approaches with a multi-axes prediction fusing have been developed[1] and compared to other network architectures.[2] For non-expert users at PETRA III, we have implemented a guided interactive and iterative framework using an active learning approach with a human in the loop in order to minimize the amount of annotations required for the training of neural networks for semantic segmentation.[3] For image enhancement, supervised and unsupervised denoising methods using a mixed-scale dense network architecture have been shown to be able to efficiently remove noise from nano- and microtomography data and outperform conventional filtering techniques.[4–6] For propagation-based phase contrast imaging, we have developed a self-supervised physics-informed learning approach for single-distance phase retrieval.[7] In

Further author information:
Julian Moosmann.: E-mail: julian.moosmann@hereon.de, Telephone: +49 (0) 4152 871048
Philipp Heuser: E-mail: philipp.heuser@desy.de, Telephone: +49 (0) 40 8998 4622

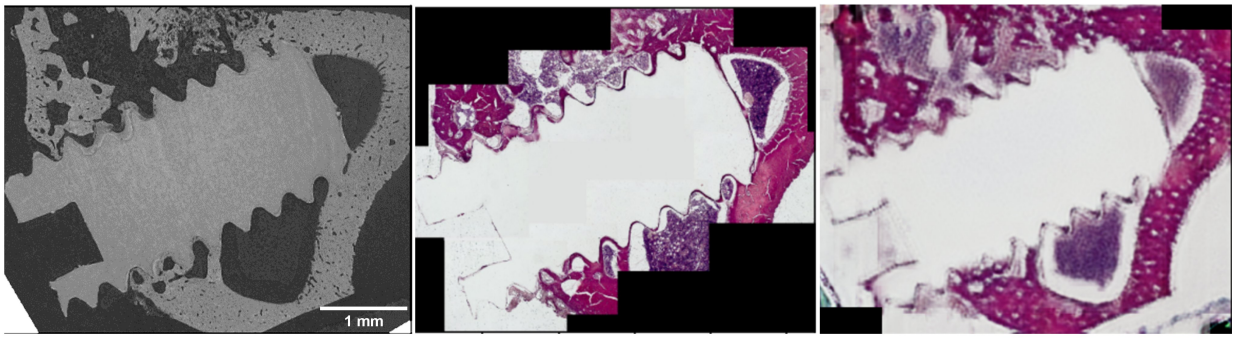

Figure 1. Generation of a virtual histology using style transfer learning. A modified cycle generative adversarial network (cycleGAN) has been trained on pairs of histological slices and corresponding registered slices of the 3D SR-µCT data. The trained network was then used to predict a virtual histology for the entire 3D SR-µCT volume. The histology/SR-µCT image pair (left/middle) was not used for training.

the following, we report on supervised and unsupervised learning for instance, see Sec. 4.1, and semantic segmentation, see Sec. 4.2, respectively. The latter can be used as a starting point for manual annotations that are then refined using the active learning framework. We further report on supervised learning approaches for the digital volume correlation of in situ, time-lapse 4D tomography data, see Sec. 3, and for the multimodal data analysis of correlative histology and tomography measurements, see Sec. 2 and Ref. 8

2. MULTIMODAL DATA ANALYSIS

In addition to synchrotron-radiation micro- and nanotomography, users at the tomography end stations at P05 and P07 regularly perform correlative characterization studies involving, e.g., histology, scanning- or transmission-electron microscopy. In contrast to the non-invasive and 3D nature of X-ray tomography, these techniques often require laborious sample preparation steps and are typically restricted to a few 2D slices, small slabs, or regions of interest. Histology, e.g., is still the gold standard for studying biological tissues but requires multiple time-consuming sample preparation steps, such as fixation, embedding, sectioning, and staining. A virtual 3D multimodal data set can be generated by extending the limited correlative 2D information to the entire 3D tomographic volume. In Ref. 9, virtual histology of soft tissues was performed by colorizing local X-ray absorption values of the 3D tomographic volume according to the histological staining. In Ref. 8, we have investigated machine learning approaches for style transfer in order to generate a 3D virtual histology of biodegradable bone implants. After registration of the histological 2D slices and the corresponding cross sections within the 3D synchrotron-radiation microtomography (SR-µCT) volume, we have trained a modified cycle generative adversarial network (cycleGAN) on a limited number of paired histology and SR-µCT images. The trained cycleGAN was then used to predict a virtual histology for the entire 3D SR-µCT volume. Preliminary results are shown in Fig. 1 with more details presented in Ref. 8. Only eight image pairs were used for training, which probably explains the blurriness of the predicted histology. The SR-µCT/histology image pair (left/middle) in Fig. 1 was not used for training. Note that the samples were critically point dried before SR-µCT scanning and subsequent histology. Therefore, the signal in the SR-µCT-tomogram is weak. In this investigation, however, our focus was not directed towards soft tissues; instead, we concentrated solely on bone tissue and its reaction to the biodegradable implant.

3. DIGITAL VOLUME CORRELATION

The high photon flux at PETRA III enables *in situ/operando/in vivo* experiments involving sample environments such as a load frame for tensile/compression/shear testing,[10] a furnace,[11] or a corrosion cell.[12,13] To evaluate the acquired 4D tomographic time-lapse sequences, we employ digital volume correlation (DVC) analysis to track deformations and strains in order to elucidate morphology-function relationships.[6] Optical flow-based DVC is

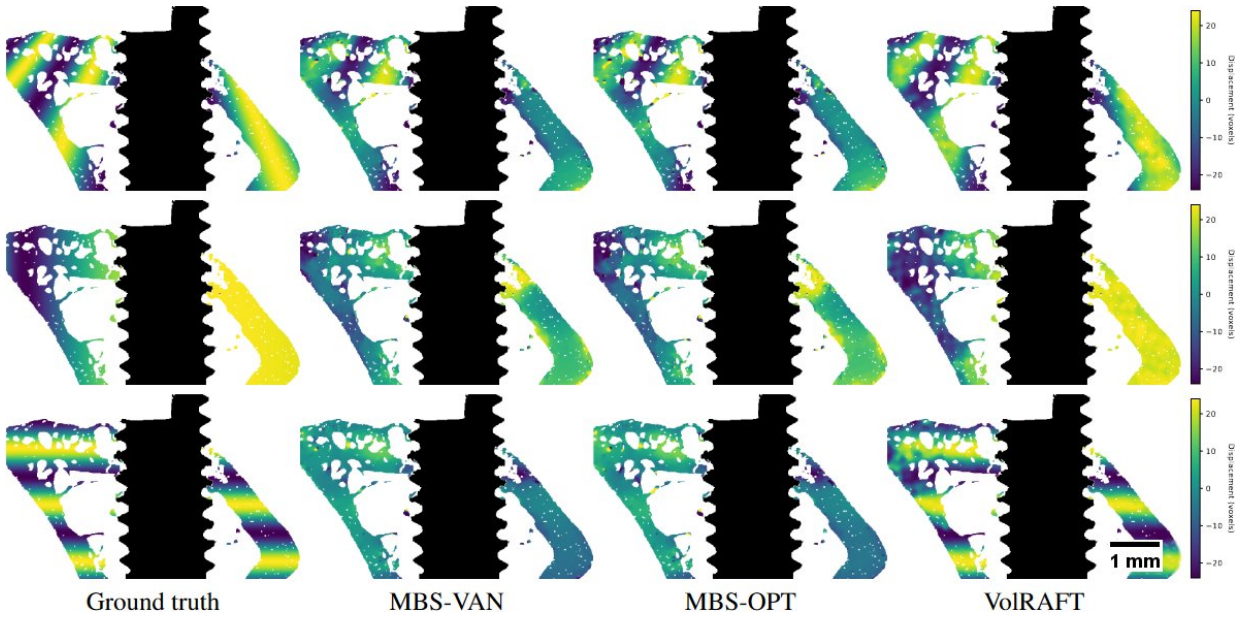

Figure 2. Digital volume correlation of 4D synchrotron-radiation microtomography data of a bone-implant interface using a volumetric optical flow network (VolRAFT). Comparison of the ground truth (first column) and displacement fields estimated by classical DVC using default (second column: MBS-VAN) and optimized settings (third column: MBS-OPT), and our learning-based approach (last column: VolRAFT). The three components of a cross section of the 3D displacement field are shown in the top, middle, bottom row, respectively. Image taken from Ref. 15. Reprinted with permission from IEEE Proceedings.

a robust method to estimate the correlation as a dense deformation vector. Recent research in computer vision showed that neural network-based optical flow approaches can outperform classical iterative optical flow-based DVC. We have extended the recurrent all-pairs field transforms (RAFT)[14] from pairs of 2D images to 3D volumes. The approach, called volumetric RAFT (VolRAFT), estimates the 3D displacement vectors between a reference and a deformed volume.[15] *In silico* experiments based on high-resolution synchrotron-radiation microtomography data of bone-implants showed that VolRAFT outperforms state-of-the-art iterative DVC methods[6] when estimating different synthetic displacement fields. Figure 2 compares the ground truth displacement fields and the fields estimated by classical methods (MBS-VAN, MBS-OPT) and our network-based approach (VolRAFT) using synthetic datasets generated from experimental data with a star-shaped displacement field. MBS-VAN denotes the vanilla version of the classical optical flow-based DVC method (MBS-3D-Optflow) based using default settings, and MBS-OPT the empirically optimized version. The source code of MBS-3D-Optflow and VolRAFT and the trained network are available on GitHub `https://github.com/hereon-mbs`.

4. SEGMENTATION

Segmentation is recurring task when analysing tomography data. In particular, the annotation of the large tomograms that are acquired during synchrotron-radiation micro- and nanotomography is often very laborious and time-consuming. In order to facilitate the segmentation of such challenging data, we have developed and implemented various machine learning approaches.[1,3] Here, we report on the task of instance segmentation and approaches for unsupervised semantic segmentation.

4.1 Instance Segmentation

Instance segmentation is the task of separating multiple objects (instances) of the same type (class) within a single volume. This is exceptionally challenging for the analysis of large volumetric SRμCT data. The example presented here is a dataset of paper fibers acquired at the Hereon imaging beamline (IBL) P05 at PETRA III at DESY,[16] where each individual fiber should be separated from the others. A very promising candidate for such

a task is EmbedSeg,[17] which already has been shown to work on the differentiation of elongated objects, such C. elegans from the BBBC010 dataset.[18] Due to hardware limitations, it is typically not possible to process the entire SR-μCT volume using 3D kernels. We thus have implemented a tile and stitch pipeline addressing the complicated joining of the obtained tile-wise segmentation results of these long fibers, which typically stretch over many tiles and may even occur more than once within one tile.[19] Ten volumes were manually annotated by a commercial partner. However, the manual annotation of such complex data is so challenging that barely sufficient annotated data was available for training and numeric evaluation. The visual evaluation of the predictions on the test case shows that the outlines of the fibers were very precisely found and within the individual tiles the individual instances were mostly well separated. However, in regions where multiple fibers are in contact, the trained network often encountered difficulties in adequately distinguishing them, so that the stitching of the individual tiles leads to the assignment of the same label to two and more fibers, see Fig. 3.

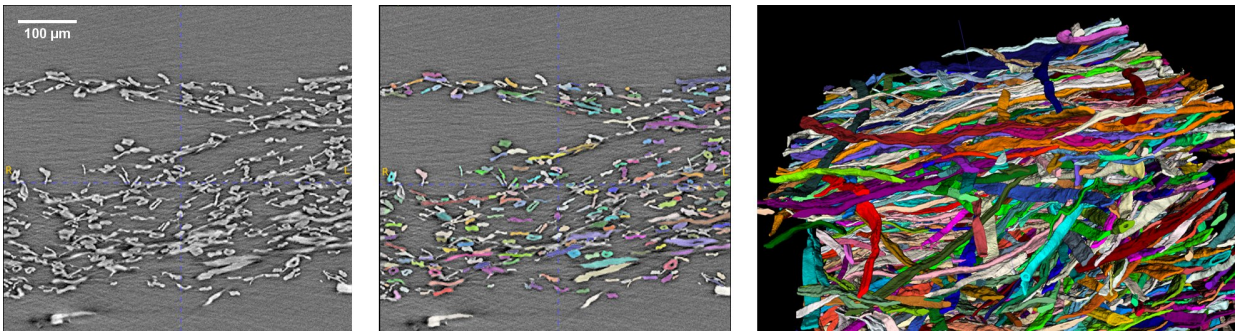

Figure 3. Instance segmentation of a tomographic data set of paper fibers. Left: Slice through the reconstructed volume showing cross-sections of individual paper fibers. Middle: Same tomographic cross-section overlayed with the instance segmentation of the individual fibers as obtained by EmbedSeg. Right: Rendering of the segmented volume showing individual paper fibers in 3D. The data was collected at P05 at PETRA III by Artem Kulachenko from the KTH Royal Institute of Technology.[16]

4.2 Unsupervised Segmentation

Inspired by the impressive results, which have been obtained for the unsupervised semantic segmentation of images from the COCO-Stuff[20] and Cityscapes[21] datasets by a self-supervised transformer with energy-based graph optimization (STEGO),[22] its application to high-resolution data from SR-μCT seems like a promising way to overcome the necessity of the tedious annotation of training data for supervised approaches. In short, STEGO exploits the already dense and semantically consistent features produced by unsupervised architectures (e.g. transformers), and optimises them to form distinct clusters. A light-weight segmentation head is added to the chosen backbone architecture to transform the features into meaningful segmentation maps. The segmentation head consists of a feedforward neural network to refine features obtained from the backbone, a clustering function to combine these features into labels, and a fully connected conditional random field optimisation to 'sharpen' these labels and improve the spatial resolution further by aligning the predictions to the image edges.[22]

STEGO is shipped with a number of pre-trained models, which are not capable to yield any usable results on SRμCT data. For example, when applied to SR-μCT data of biodegradable bone implants, a mean intersection over union (IoU) of around 0.21-0.24 is obtained, see Fig. 4. However, subsequent to the retraining of the segmentation head utilizing images from the dataset, some of the features of the samples become visible in the prediction of the unsupervised segmentation. After retraining, the mean IoU of the prediction is in the order of 0.5. While the bone and screw can be identified, the precision with which their boundaries are identified is not yet sufficient. Identifying those pixels which represent the degraded screw is not possible, see Fig. 5.

5. DISCUSSION AND CONCLUSION

The preliminary results presented for the virtual histology are very promising considering the limited amount of training data used. At present, we are expanding the dataset and intend to incorporate histological data stained

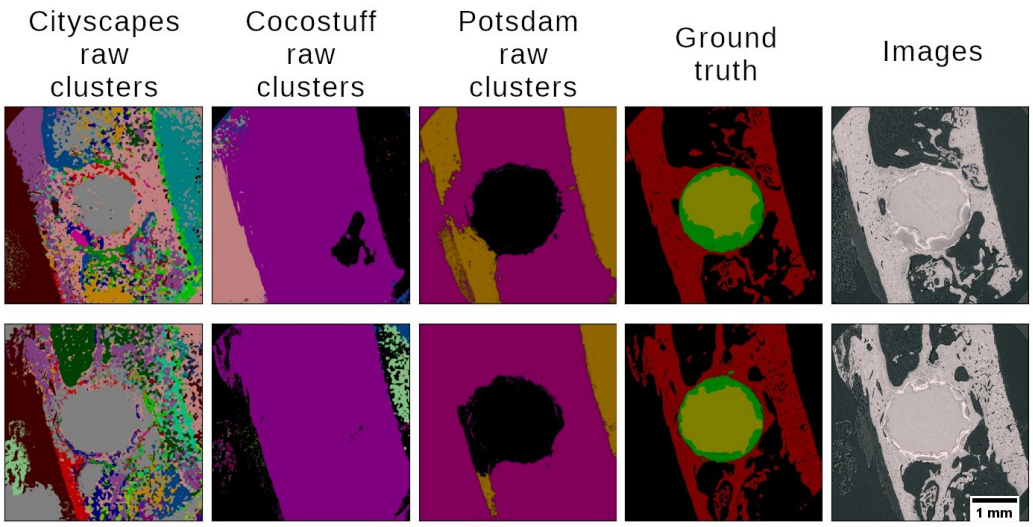

Figure 4. Unsupervised segmentation (without retrained segmentation head). Application of the models trained on the respective datasets to 2D slices of the SR-µCT data of biodegradable bone implants.

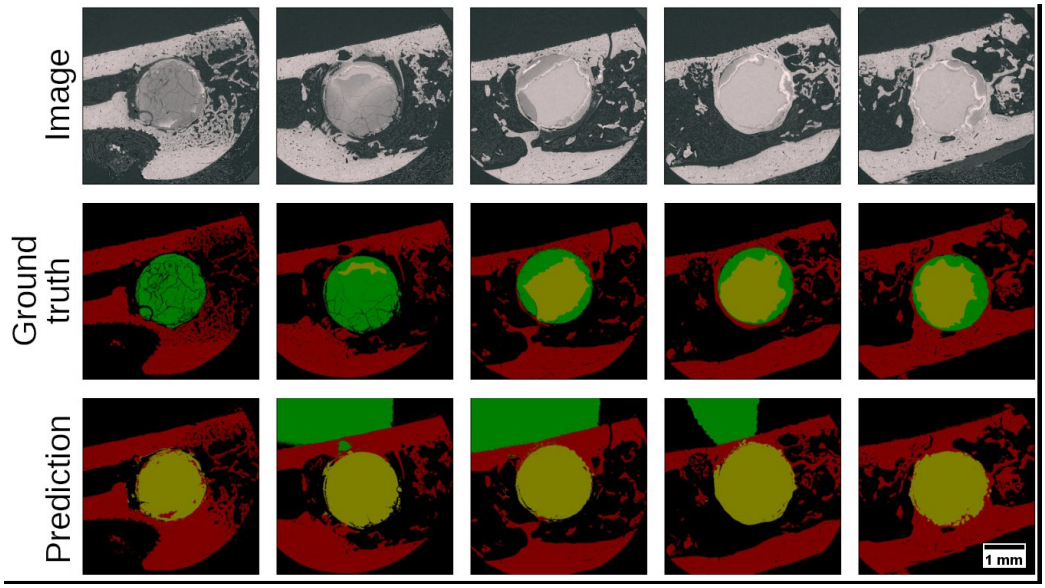

Figure 5. Unsupervised segmentation with retrained segmentation head. Application of STEGO models after retraining of the segmentation head with images from the SR-µCT dataset and their application to test data.

with toluidine blue. This could potentially improve resolution and contrast. We further aim to optimize the network hyperparameters and explore alternative variants of GAN-based networks. Ultimately, our goal is to establish a digital pathology that preserves the specificity of biological staining throughout the 3D volume while minimizing experimental effort.

A learning-based methodology for optical flow-based digital volume correlation has been effectively extended from 2D images to 3D tomograms. This approach (VolRAFT) does not require further fine-tuning or optimization of network hyperparameters. *In silico* experiments using synthetic displacement fields have demonstrated that VolRAFT surpasses cutting-edge iterative methods. Applied on measured (non-synthetic) datasets, it further shows great generalizability, as it can predict the overall structure of the displacement field even if the testing set is unknown in the training set of the VolRAFT model. In subsequent research, we intend to assess the

performance of VolRAFT on additional datasets and to enhance the categories of synthetic displacement fields employed for training. This will be achieved by conducting a principal component analysis of displacement fields that have been classically computed from measured 4D data.

Machine learning has been effectively utilized for the instance segmentation of fiber bundles. Considering the spatially curled orientation of these fibers, it is imperative to employ a three-dimensional neural network. Due to inherent memory constraints, a patch-based methodology is required, followed by subsequent stitching. Although this tile-and-stitch strategy is fundamentally effective, it requires annotated training data, whose procurement is labor intensive and costly. As the subsequent stitching occasionally leads to ambiguities, we are in the process of optimizing the respective pipeline to separate the individual fibers.

We further explored approaches for unsupervised segmentation using STEGO. The immediate deployment of pre-trained models did not yield satisfactory outcomes. However, retraining the segmentation head significantly improved predictive performance, despite the inability to identify all relevant labels. It should be acknowledged that the data set in question presents a significant challenge for segmentation as the corrosion layer of the biodegradable implant is characterized by a highly fragmented and variable texture.

ACKNOWLEDGMENTS

We thank Dr. Stefan Bruns for his valuable insight and the fruitful discussions regarding the optimization of the MBS-Optflow framework. Parts of this research were supported by the BMBF project 'Multi-task Deep Learning for Large-scale Multimodal Biomedical Image Analysis (MDLMA)' (project number 031L0202A), the Hereon project 'Holistic Data Analysis (HoliDAy)' of the Innovation-, Information-& Biologisation-Fonds ((I^2B)), the ErUM-Data Verbundprojekt 'KI4D4E: Ein KI-basiertes Framework für die Visualisierung und Auswertung der massiven Datenmengen der 4D-Tomographie für Endanwender von Beamlines' (project number 05D23CG1) which is funded by the Bundesministeriums für Bildung und Forschung (BMBF), HELMHOLTZ IMAGING, a platform of the Helmholtz Information & Data Science framework, and the Helmholtz Joint Laboratory 'Model and Data-driven Materials Characterisation (JL MDMC)', a cross-centre platform of the Helmholtz Association (HGF). We acknowledge the Deutsches Elektronen-Synchrotron DESY (Hamburg, Germany), a member of the HGF, for the provision of beamtime at the imaging beamline (IBL) P05 and high energy materials science beamline (HEMS) P07 at PETRA III at DESY related to the proposals I-20160104, II-20170009, I-20170074, I-20180109, I-20180842, I-20191467 and I-20190205. This research was supported in part through the Maxwell computational resources operated at DESY and the Trilateral Data Science Exchange Program, hosted by the British Council, the Israel Data Science and AI Initiative (IDSI) and the Helmholtz Information & Data Science Academy.

REFERENCES

[1] Baltruschat, I. M., Cwieka, H., Krüger, D., Zeller-Plumhoff, B., Schlünzen, F., Willumeit-Römer, R., Moosmann, J., and Heuser, P., "Scaling the u-net: Segmentation of biodegradable bone implants in high-resolution synchrotron microtomograms," *Scientific Reports* **11**(24237), 1–10 (2021).

[2] Lopes Marinho, A., Kazimi, B., Ćwieka, H., Marek, R., Beckmann, F., Willumeit-Römer, R., Moosmann, J., and Zeller-Plumhoff, B., "A comparison of deep learning segmentation models for synchrotron radiation based tomograms of biodegradable bone implants," *Frontiers in Physics* **12** (2024).

[3] Kazimi, B., Heuser, P., Schluenzen, F., Cwieka, H., Krüger, D., Zeller-Plumhoff, B., Wieland, F., Hammel, J. U., Beckmann, F., and Moosmann, J. P., "An active learning approach for the interactive and guided segmentation of tomography data," in [*Developments in X-Ray Tomography XIV*], Müller, B. and Wang, G., eds., SPIE (nov 2022).

[4] Flenner, S., Storm, M., Kubec, A., Longo, E., Döring, F., Pelt, D. M., David, C., Müller, M., and Greving, I., "Pushing the temporal resolution in absorption and Zernike phase contrast nanotomography: enabling fast *in situ* experiments," *Journal of Synchrotron Radiation* **27**, 1339–1346 (Sep 2020).

[5] Flenner, S., Bruns, S., Longo, E., Parnell, A. J., Stockhausen, K. E., Müller, M., and Greving, I., "Machine learning denoising of high-resolution X-ray nanotomography data," *Journal of Synchrotron Radiation* **29**, 230–238 (Jan 2022).

[6] Bruns, S., Krüger, D., Galli, S., Wieland, D. F., Hammel, J. U., Beckmann, F., Wennerberg, A., Willumeit-Römer, R., Zeller-Plumhoff, B., and Moosmann, J., "On the material dependency of peri-implant morphology and stability in healing bone," *Bioactive Materials* **28**, 155–166 (2023).

[7] Yang, X., Hailu, D., Kulvait, V., Jentschke, T., Wong, T. M., Flenner, S., Greving, I., Hagemann, J., Campbell, S., Schroer, C., and Moosmann, J., "Physics-informed generative network for phase retrieval from a single intensity measurement," (2024).

[8] Irvine, S., Lucas, C., Bootbool, M., Galli, S., Zeller-Plumhoff, B., and Moosmann, J., "Multi-modal image registration and machine learning for the generation of 3d virtual histology of bone implants," in [*Developments in X-Ray Tomography XV*], Müller, B. and Wang, G., eds., **13152**(13152-70), SPIE (2024).

[9] Khimchenko, A., Deyhle, H., Schulz, G., Schweighauser, G., Hench, J., Chicherova, N., Bikis, C., Hieber, S. E., and Müller, B., "Extending two-dimensional histology into the third dimension through conventional micro computed tomography," *NeuroImage* **139**, 26–36 (2016).

[10] Moosmann, J., Wieland, D. C. F., Zeller-Plumhoff, B., Galli, S., Krüger, D., Ershov, A., Lautner, S., Sartori, J., Dean, M., Köhring, S., Burmester, H., Dose, T., Peruzzi, N., Wennerberg, A., Willumeit-Römer, R., Wilde, F., Heuser, P., Hammel, J. U., and Beckmann, F., "A load frame for in situ tomography at PETRA III," in [*Developments in X-Ray Tomography XII*], Müller, B. and Wang, G., eds., **11113**, 1111318, International Society for Optics and Photonics, SPIE (2019).

[11] Tolnai, D., Gavras, S., Wilde, F., Hammel, J. U., and Bruns, S., "In situ synchrotron tomography of the solidification of an elektron 21 mg alloy," *Advanced Engineering Materials* **23**(11), 2100383 (2021).

[12] Zeller-Plumhoff, B., Helmholz, H., Feyerabend, F., Dose, T., Wilde, F., Hipp, A., Beckmann, F., Willumeit-Römer, R., and Hammel, J. U., "Quantitative characterization of degradation processes in situ by means of a bioreactor coupled flow chamber under physiological conditions using time-lapse srÂµct," *Materials and Corrosion* **69**(3), 298–306 (2018).

[13] Reimers, J., Trinh, H. C., Wiese, B., Meyer, S., Brehling, J., Flenner, S., Hagemann, J., Kruth, M., Kibkalo, L., Ćwieka, H., Hindenlang, B., Lipinska-Chwalek, M., Mayer, J., Willumeit-Römer, R., Greving, I., and Zeller-Plumhoff, B., "Development of a bioreactor-coupled flow-cell setup for 3d in situ nanotomography of mg alloy biodegradation," *ACS Applied Materials & Interfaces* **15**(29), 35600–35610 (2023). PMID: 37459562.

[14] Teed, Z. and Deng, J., "Raft: Recurrent all-pairs field transforms for optical flow," in [*Computer Vision – ECCV 2020*], Vedaldi, A., Bischof, H., Brox, T., and Frahm, J.-M., eds., 402–419, Springer International Publishing, Cham (2020).

[15] Wong, T. M., Moosmann, J., and Zeller-Plumhoff, B., "Volraft: Volumetric optical flow network for digital volume correlation of synchrotron radiation-based micro-ct images of bone-implant interfaces," in [*2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*], 53–62 (2024).

[16] Wallmeier, M., Barbier, C., Beckmann, F., Brandberg, A., Holmqvist, C., Kulachenko, A., Moosmann, J., Östlund, S., and Pettersson, T., "Phenomenological analysis of constrained in-plane compression of paperboard using micro-computed tomography imaging," *Nordic Pulp & Paper Research Journal* **36**(3), 491–502 (2021).

[17] Lalit, M., Tomancak, P., and Jug, F., "Embedding-based instance segmentation in microscopy," (2021).

[18] Ljosa, V., Sokolnicki, K. L., and Carpenter, A. E., "Annotated high-throughput microscopy image sets for validation," *Nature Methods* **9**, 637–637 (July 2012).

[19] Heuser, P., "Helmholtz imaging patch generation and merging tool - hi patch," (2022).

[20] Caesar, H., Uijlings, J., and Ferrari, V., "Coco-stuff: Thing and stuff classes in context," in [*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 1209–1218 (2018).

[21] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., and Schiele, B., "The cityscapes dataset for semantic urban scene understanding," in [*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], 3213–3223 (2016).

[22] Hamilton, M., Zhang, Z., Hariharan, B., Snavely, N., and Freeman, W. T., "Unsupervised semantic segmentation by distilling feature correspondences," (2022).