PROJECT ACRONYM:

MDLMA

PROJECT TITLE:

Multi-task Deep Learning for Large-scale Multimodal Biomedical Image Analysis

PROPOSER:

Prof. Dr. Regine Willumeit-Römer
Dr. Julian Moosmann
Institute of Materials Research
Helmholtz-Zentrum Geesthacht (HZG)
Zentrum für Material- und Küstenforschung
Max-Planck-Str.1
21502 Geesthacht
Tel: +49 40 8998-6904
Mobile: +49 170 4316430
Email: regine.willumeit@hzg.de

IN COOPERATION WITH:

Prof. Dr. Mattias Heinrich
Institute for Medical Informatics (IMI)*

Prof. Dr. Thorsten M. Buzug
Institute for Medical Engineering (IMT)*

*Universität zu Lübeck (UzL)

AND:

Dr. Frank Schlünzen
Dr. Philipp Heuser, Dr. Xiaogang Yang
Deutsches Elektronen-Synchrotron DESY

ASSOCIATED PARTNER:

Dr. Jan-Marten Seitz
Syntellix AG

Total costs: 1,088,328 EUR

Total funding volume: 1,170,157 EUR

Project duration: 36 month

1 Aims of the project

It is a common scientific approach to investigate a physical object with multiple, complementary experimental methods. In this project we are focusing on certain classes of biomaterials which are characterized using laboratory x-ray computed tomography (labCT), synchrotron radiation micro-computed tomography (SRμCT), magnetic resonance imaging (MRI), small angle x-ray scattering (SAXS) and histology. Experimental imaging data consequently comprises varying spatiotemporal resolution, dimensionality and modality-specific appearance, which is a very common problem encountered in many data-driven scientific fields. Analyzing each of the datasets and extracting relevant information individually, such as object segmentations or noise reduction, is already challenging, due to the complexity and large scale of the acquired data. Combining the output of several separately processed datasets, e.g. images of different modalities or segmentations and image enhancement results, becomes even more difficult and error-prone due to the lack of analytically determinable transformations that would restore the relation of partially overlapping object characterizations.

Within this project, we first aim to address individually the most important tasks of image analysis including registration, segmentation/classification, and artifact or noise reduction using deep learning (DL) approaches for the specific biomedical task of bone implant characterization. Second, we will devise new multi-task deep learning methods that are able to integrally combine different complementary tasks and transfer the knowledge across analysis tasks. Therefore, several joint models for simultaneous registration and segmentation or image enhancement with segmentation priors can be realized. Third, we aim to identify common feature representations that are learned within generic structures of deep networks for various image analysis tasks, such as shape and appearance priors. By explicitly decoupling the task-specific network parts, we aim to find a unified data-driven approach that will enable its application to a much broader scope of domains, modalities, artifacts, resolutions, etc. Once a generic network architecture is developed with the help of domain experts, the goal is that further expert knowledge for problems in similar domains is not required anymore. Such problems are very common in Life Sciences (and other fields) ranging from analyses of anatomical changes over time in longitudinal brain MRI studies to evaluations of the performance of biodegradable implants in animal studies or humans to multimodal multiresolution integrative macromolecular structural biology experiments.

2 Scientific concept

The specific problems of registration, segmentation and artifact reduction or image enhancement will be first addressed individually using deep learning and then using multitask-learning approaches in a combined fashion at various levels and stages with a recurrent or iterative scheme. This will be demonstrated exemplarily on research on biodegradable implants since a plethora of multimodal data from multiple methods and devices at various resolutions is already available or currently collected by HZG and Syntellix (see Figure 1 and Chapter 4).

2.1 Scientific objectives

The aim of this project is to develop new methods for deep learning in biomedical imaging that leverage large datasets acquired with different modalities and by employing multiple complementary supervised tasks (reconstruction, segmentation, and registration) within a single network to improve the learned image understanding. In contrast to the vast majority of currently employed approaches that train a specialized deep learning model for each modality and each analysis task separately, we aim to disentangle the input data into a representation of its underlying semantic components (e.g. shape, object parts) and the given modality and scanner specific appearance. These powerful priors on image shape and appearance will be subsequently used to substantially improve the accuracy of individual analysis tasks e.g. the iterative reconstruction of tomographic images, the segmentation of challenging bone implant data, and the multimodal registration of MRI, labCT, SRμCT and histology.

In the following, the principal scientific research questions will be presented in detail.

2.1.1 Scientific background of the data collected

Permanent implants made of Titanium (Ti) or its alloys are the gold standard in most orthopedic and traumatological applications. However, in the case of children still growing or on patient's demand a follow-up surgical intervention is required to remove the implant which is an additional health risk and cost factor (costs for explantation without complications in 2010 in Germany: 281 M€). Moreover, detrimental effects can be caused by stress shielding, wear debris or allergic reactions. To avoid such complications biodegradable implants are clearly preferred [1]. A promising implant material is Magnesium (Mg) due its biocompatibility (Mg is an essential element required by the human body), mechanical stability (Mg implants and bone have similar Young's modulus), and degradation properties (Mg degrades under physiological conditions into products well tolerated by the human body). However, a fast and uncontrolled corrosion results in a strong release of hydrogen and ions and severe pH changes which can cause a loss of mechanical stability of the implant and undesirable biological reactions [2]. Using alloying elements the degradation and mechanical properties of the implant can be tailored such that the bone defect is stabilized during healing and implant degradation. These are highly complex processes in the living system and sufficient in vivo data is missing [3]. Therefore, the division *Metallic Biomaterials* at HZG is conducting extensive research to gain a most comprehensive characterization of the bone-implant interface. In order to establish Mg as implant material, the following questions will be addressed:

- What are the underlying mechanisms of the in vivo degradation of implants in bone?
- How does the implant degradation influence the encapsulation of the implant by bone (osseointegration)?
- How do different implant materials stimulate bone formation at the bone-implant interface (bone remodeling)?
- How is stress dissipated at the interface and what are the failure mechanisms under mechanical load?

Outcome: An improved understanding of the degradation, osseointegration, bone remodeling and failure mechanisms will enable the design of implants that are optimized for medical applications.

Addressed funding priorities: A plethora of multimodal data is currently collected by HZG in order to characterize the bone-implant interface. Since data processing and analysis is challenging due to the amount and heterogeneity of the collected data sets, advanced multi-task and multimodal analysis tools will be developed within this project.

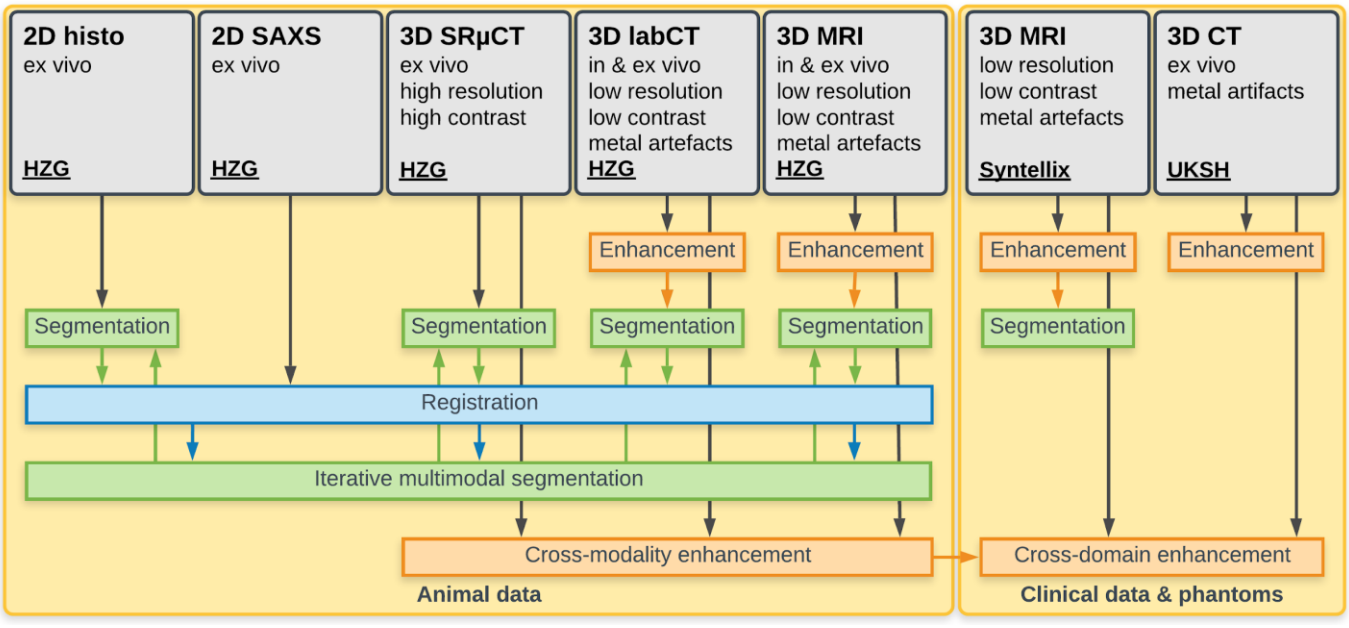

Figure 1: Schematic overview of imaging modalities, data sets and tasks to be performed.

2.1.2 Image enhancement

Our central hypothesis is that a complementary combination of deep learning and mathematical modelling will lead to improvements in iterative CT reconstruction and multimodal artifact reduction. The following questions will be investigated in this project:

- How much can the image quality be improved by post-processing with convolutional neural networks (CNNs) and is it possible to transfer already learned parameters to different image modalities?
- Can a learned adversarial loss function penalize artifacts in CTs during the reconstruction?
- Does the joint integration of an iterative reconstruction algorithm with (adversarial) CNNs offer further advantages?
- Does learning a sparse system matrix provide further runtime and robustness benefits?
- Are learned and non-local priors advantageous for weighted image regularization?
- How can the transfer of the learned methodology, especially the multitask approach, to phantom and human CT data be achieved?

Outcome: It is to be expected that new insights will be gained regarding the integration of different modalities to achieve an improved image quality. In particular, the combination of deep learning and mathematical modelling will lead to better image quality in terms of recognition of anatomical details, the amount of artifacts and basic prerequisite for further processing.

Addressed funding priorities: The project addresses the application of innovative and experimental concepts for the integration and combination of various image data (e.g. learned prior based on different imaging modalities) based on deep learning techniques combined with conventional analytical and numerical methods.

2.1.3 Segmentation

We propose that the use of multiple datasets of the same sample collected with different experimental methods yielding images at varying resolutions, with varying level of detail and different appearance (e.g. histology) improve the accuracy of the segmentation for various image modalities. Following questions will be addressed:

- Investigation of segmentation-accuracy improvement due to the simultaneous usage of multimodal data for segmentation.
- Possibilities of cross-modality transfer learning. Do the segmentations of the individual data types benefit from networks trained on multimodal data? As a result, imaging using lower resolution and lower dose experiments can provide results comparable to those obtained today only with SRμCT, for example, making the expensive synchrotron experiments obsolete.
- Is it possible to learn rotation and scale invariant or covariant filter kernels by using deformable convolution training that enables an adaptive geometric transformation of the spatial filter layouts.

Outcome: Segmentations for different biomedical data types used as priors for registration and enhancement, and subsequently a segmentation tool capable of using multiple data types in parallel with emphasis on precision and exploiting the heterogeneity of different experimental data. The aim is to develop a model that can use the accumulated knowledge from all modalities to achieve a segmentation of the individual data types which is superior to the one using a single data type only.

Addressed funding priorities: We address the influence of weakly annotated heterogeneous data (i.e. incomplete, insufficient and inadequate manual segmentations from different sources and modalities, which further depend on the experience of the domain expert) with respect to the reliability and quantitative interpretability of a segmentation tool.

2.1.4 Registration

We will develop novel deep learning based algorithms for accurate and robust multimodal image registration. Instead of directly learning a very complex network that directly regresses a (non-linear) transformation given two inputs, we aim to decouple the feature learning part of the network and separate it from a mathematically motivated geometric alignment part to answer the following research questions:

- Is it possible to train a deep multimodal registration network from much fewer labelled training data by explicitly using iterative geometric transformation models?
- Can the robustness and accuracy be further improved by learning a prior model that captures anatomical information independent of modality to provide a semantic guidance for the alignment task?
- Will transfer and/or multi-task learning further assist the convergence of image registration by initializing network weights based on segmentation and image reconstruction models?
- Is it possible to model the iterative fitting of spatial transformations in a differentiable manner to train a multi-step deep alignment model in an end-to-end fashion and thereby enable a robust coarse-to-fine registration?

Outcome: We expect to develop more robust deep registration models that have fewer trainable weights and can be trained with weak labels and/or intrinsic information from large datasets. We will reduce the computational burden of large-scale registration tasks by designing efficient multi-stage feed forward transformer networks. The software modules will be beneficial in other work tasks by improving transfer learning through multimodal fusion of information.

Addressed funding priorities: The work will incorporate existing domain knowledge (iterative mathematically modelled geometric alignment) into the deep models. Our multimodal alignment will connect different imaging modalities to capture the complex interactions in this data and enable transfer learning for other analysis tasks.

2.1.5 Multi-task learning

We will design a concept for multi-task learning targeted at multiscale and multimodal biomedical data that comprises of a foundational prior network and specialized task- and modality-specific network parts that are jointly trained. The following research questions will be addressed in this work:

- How can we best design a common foundational network architecture that captures normal anatomical priors and their variations of the studied objects with high robustness?
- Will this generalistic network enable us to effectively study abnormalities (pathologies, temporal changes of material composition, etc.) by training several task specific sub-networks?
- Which recently developed techniques for multi-task learning beyond simple parameter-sharing such as cross-stitch networks [4], guidance through attention-gate networks [5], etc. are best suited for multimodal biomedical image analysis? How can these be developed further and is it potentially possible to automatically search the best combinations of those [6]?

Outcome: Development of appropriate methods pre-trained networks for deep multi-task learning adapted for biomedical data challenges which capture priors of considered anatomical sites and modalities. Evaluation of the ability to transfer knowledge across multiple biomedical image analysis tasks.

Addressed funding priorities: Increased interpretability and robustness of models by learning explicit prior models of shape. Reducing the reliance on training models from scratch with specific expert domain knowledge by adapting common pre-trained architectures that transfer information and experience.

2.1.6 Unified framework

We propose to implement the methods and tools developed for image enhancement, segmentation and registration in a unified framework to address these tasks simultaneously across domains and modalities using multi-task learning strategies. The following questions will be addressed:

- What are the most important generic building blocks of each task considered and how can they be effectively shared within a unified framework?
- How can a simultaneous multi-task data analysis potentially help to improve the domain adaptation problem and increase the accuracy and robustness of individual tasks?
- Is it possible to improve the generalization of image analysis models through scale-adaptive and modality-invariant filter kernels to enable a more effective transfer learning?
- How can additional tasks be implemented in a unified framework and trained with little expert supervision to support an easier adaptation of complex deep networks in life sciences?

Outcome: A unified framework for multi-task analysis will foster synergies and is expected to be highly beneficial for a broad user community because the tasks considered are relevant to most users working with imaging data.

Addressed funding priorities: For the implementation of the framework free and open-source machine learning libraries will be used such as TensorFlow or PyTorch. The performance of the framework will be evaluated across domains and modalities.

2.2 Methodological approach

2.2.1 Imaging modalities and data analysis

In order to elucidate the degradation behavior, osseointegration and failure mechanisms of biodegradable implants, HZG aims at most comprehensive characterization of the bone-implant interface using a multimodal, multiscale approach including labCT, SRµCT, SAXS, MRI and histology. Several different Mg alloys are considered as implant materials as well as Polyetheretherketone (PEEK) and Titanium (Ti) as reference materials. The data sets collected are thus prone to different artifacts depending on the implant material as well as the imaging modality e.g. low contrast of Mg alloys and PEEK in labCT or low contrast and metal artifacts of Ti and Mg alloys in MRI and labCT (see Figure 2). Following data sets will be provided by the consortium within this project:

- Ex vivo data of implants (Mg alloys, PEEK, Ti) in rat femurs including 3D volumes from SRµCT and 2D slices from SAXS, histology. SRµCT is available for critically-point dried samples, fresh samples from biomechanical tests and fresh samples originating from previous in vivo labCT studies.
- In vivo scans of implants (Mg alloys, PEEK, Ti) in rat femurs using low-resolution MRI and labCT. Both modalities are prone to artifacts (metal, motion, low contrast, etc.) artifacts.
- At the end of the in vivo studies, ex vivo MRI and labCT will be acquired with scan parameters optimized for image quality as neither radiation dose nor movement artifact will be of concern. In addition, SRµCT will be acquired with high spatial resolution and contrast and without beam hardening artifacts due to monochromatic synchrotron radiation.
- Clinical MRI of biodegradable implants in patients provided by Syntellix AG.
- Clinical data from other areas such as patients with hip or knee implants, orthopedic screws and stents is provided by University Medical Center Schleswig-Holstein (UKSH).

Figure 1 depicts a schematic overview of imaging modalities, available data sets and tasks to be performed. Figure 2 shows images acquired from different modalities and the registration of SRµCT, SAXS and histology.

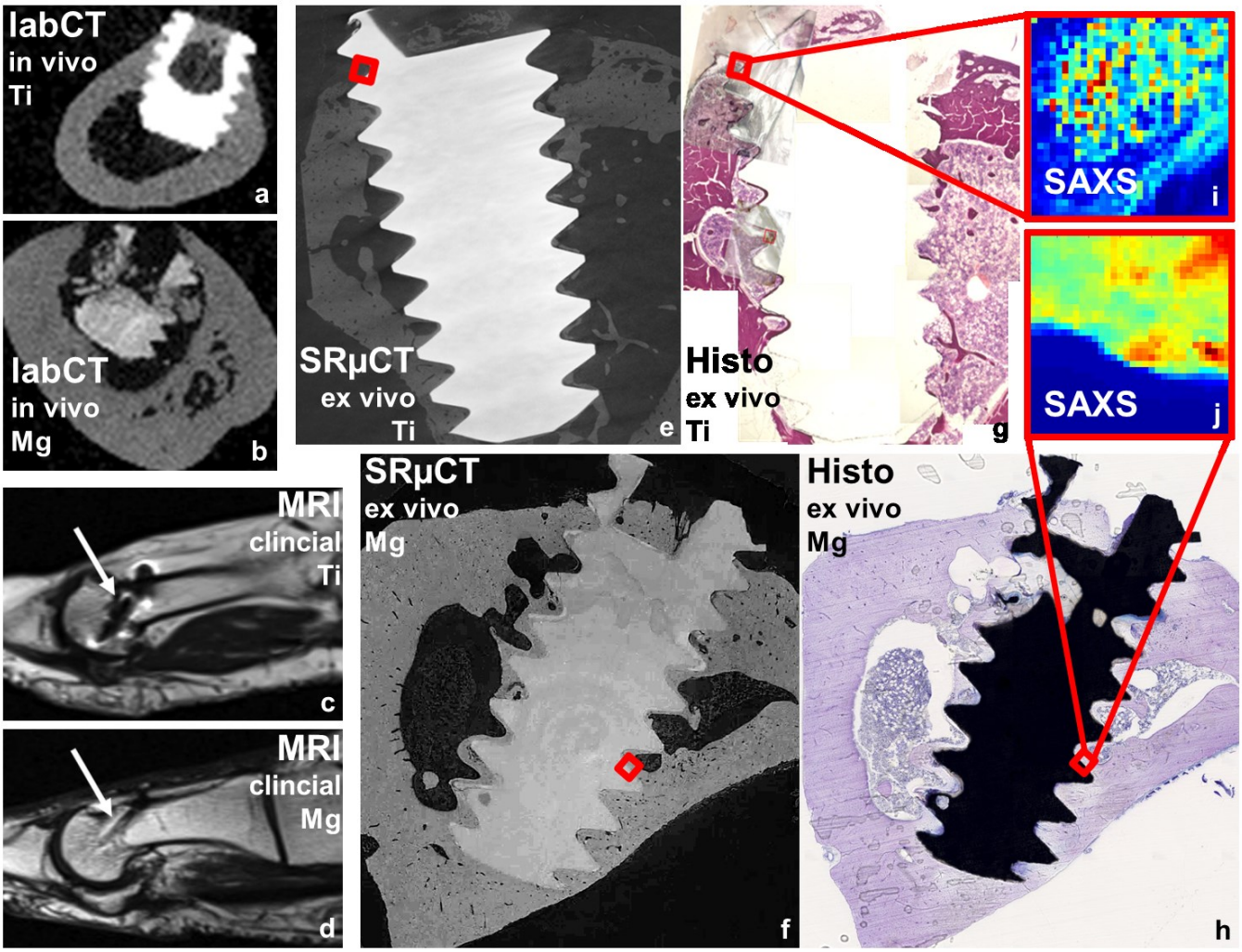

Figure 2: Imaging modalities and multimodal registration: (a,b) Animal data from in vivo labCT of a permanent Ti (a) and a biodegradable Mg (b) implant in a rat femur. Note that metal artifacts are minor in (a) due to the orientation of the depicted slice and the used window and level. (c,d) Clinical MRI of a human forefoot three years after a hallux valgus correction (osteotomy) using a Titanium (c) and Mg (d) implant, respectively. (e-j) Ex vivo animal data from SRμCT (e,f), histology (g,h) and SAXS (i,j) of the same Ti-implant (e,g,i) and Mg-implant (f,h,j), respectively. The red box in (e-h) indicates where SAXS measurements where performed. (e-h) Registration of multimodal data from SRμCT, histology and SAXS. A landmark-based method with rigid transformation was used to register the 2D slice from histology and SAXS with the 3D volume from SRμCT. Note that the histological sections in (g) are superimposed by images from optical microscopy.

2.2.2 Image enhancement

A partial application of the project is to achieve significant methodological progress in the field of CT image reconstruction taking various approaches for artifact correction into account. By combining the recently developed iterative reconstruction techniques augmented likelihood image reconstruction (ALIR) and current deep learning techniques, the state of the art for inverse problems in medical imaging will be improved, which is expected to have far-reaching implications for the whole field [7]. The focus of our planned work lies in a detailed investigation of different points of contact and mutual reinforcement between data-driven, monitored learning procedures, which independently grasp semantic correlations with the help of large data sets, and analytical procedures, which can all combined yield specific geometric and anatomical prior knowledge.

In the planned period of three years, these questions will be answered systematically and evaluated quantitatively for various challenging reconstruction problems. UzL is highly competent in generating training and validation data with realistic (software) phantoms and has access to comprehensive clinical radiological image data of human individuals. Furthermore, image data of rats from SRμCT and labCT can be used to compare Mg-based, PEEK-, and Ti-implants regarding the severity of artifacts and image

enhancement ability. On the one hand, the focus of the research will be directly on improving image quality and avoiding artifacts and, on the other hand, on the development of new methods which are suitable for further applications in inverse problems in medical imaging and image processing. Based on the previously presented ALIR algorithm, new methods for the reduction of various artifact types will be developed [7]. In particular, we will investigate how deep learning techniques can be used to generate a prior, suppress artifacts and directly reconstruct images. The base of the sub-project is a combination of robust numerical procedures with data-driven learning algorithms, whereby the complementary approaches mutually support each other. Convolutional networks can usually learn better filters for edge preservation through extensive training and reliably distinguish between disturbed and artifact-free images through discriminative loss functions. It is therefore of great interest to examine whether iterative reconstruction algorithms in combination with feed-forward networks will provide further advantages and improvements.

The ALIR algorithm is an iterative reconstruction approach in which projection values that are associated with metal are replaced in each iteration. As a new approach, the replacement of projection values will be adapted to various types of artifacts, e.g. motion artifacts, noise reduction or truncation artifacts. In each iteration, newly calculated projection values will be combined with the original measurement. For the exemplary application of noise artifacts, intermediate results of the iterative reconstruction can be filtered in order to reduce noise, or CNN-based images in combination with other modalities can be created. The basic idea here is that artificially generated projections contribute to noise reduction in the image while original measurement data ensure that specific anatomical details are correctly reconstructed. A projection replacement step based on a direct deep learning reconstruction and CNN-based filter for artifact reduction will be investigated in detail.

Furthermore, images generated by a CNN offer great potential for the application of a prior image based regularization. Images can be derived either from interim results of the reconstruction, different image modalities such as SRµCT and MRI, or a combination of the two. The use of generative adversarial networks will be investigated. Once a prior image has been defined, it is to be integrated into the reconstruction with the help of a regularization term. The aim is to use information about the anatomy of the object of investigation from the prior to reduce artifacts in the reconstructed image. An important aspect is not to use the anatomy from the prior to perform a pixel-to-pixel comparison and penalize intensity differences. This would result in not mapping the correct anatomy of the patient, but only a copy of the prior image. To avoid this effect, a non-local regularization term is used, which is presented for a software phantom in the preliminary work [8]. Such a procedure is particularly useful when using images generated by a CNN as the prior no longer has to fulfill the property that it depicts anatomically correct and positionally true structures, but should fulfill the condition with the help of a learned discriminative loss function that the depicted structures are related to the measured raw data and that the generated image is artifact-free.

In the specific case of metal artifacts, the proposed regularization will be extended by a local weighting approach. The magnitude and location of metal artifacts is strongly dependent on the shape of the metal object. Using an initially reconstructed image, which is either the result of a direct deep learning reconstruction or a CNN filtered reconstruction, one can calculate an opening angle to the metal object for each pixel in the image which indicates the magnitude of an artifact at a specific position. Furthermore, this angle can be used for a weighting scheme of the regularization.

A lot of the proposed methods are specific for CT imaging, especially deep learning based approaches that are developed for post-processing of either reconstructed images or interim results. Thus, these methods will be transferred for an application on other imaging modalities such as MRI. Image enhancement (e.g. spatial resolution, artifact reduction) of in vivo MRI/CT data from rats will be achieved by cross-domain learning from high-quality SRµCT images and the quality of longitudinal clinical MRI data will be improved by cross-domain learning from ex vivo scans.

2.2.3 Segmentation

For the detailed quantitative analysis of the degradation process of biomaterials, very precise semantic segmentation of several hundreds of collected 3D tomographic datasets is crucial. The scope of this task goes beyond the usual requirements for semantic segmentation where often the localization and rough definition of the boundaries of a region of interest is sufficient.

X-ray tomographic volumetric data of biomaterials are particularly challenging. The precise boundaries between different materials and tissues of interest are – even for trained experts – hardly to distinguish and difficult to interpret. This often induces small errors in state-of-the-art segmentation methods including small false positive patches, imprecise boundaries of objects, etc., which are irrelevant in many applications, but significantly influence the results of scientific quantitative biomedical analysis of 3D tomographic volumes where e.g. the evolution of contact areas over time are studied.

The large amount of data – a single SRμCT data set typically has about 1500^3 voxels – has made semi-automatic extensions of manually segmented slice, e.g. by interpolation with watershed algorithms, indispensable. However, such a procedure unavoidably propagates and increases errors in the initial segmentation. Particularly iterative procedures lead to a poorly reproducible segmentation impairing the quality of the expert's manual segmentation. Consequently, the annotations for deep learning based segmentations of tomographic images are often inaccurate, which bears challenges for the training and evaluation of deep CNNs. In particular, for the evaluation of the accuracy, stability and the comparability of models additional measurements need to be developed. Ultimately, removing the necessity of semi-automatic and error-prone procedures is one of the goals of the multi-task deep learning pipeline.

For the development and training of deep learning methods for precise and accurate automatic semantic segmentation and their quantitative analysis, we will study the robustness of deep learning architectures towards inaccuracies within manually annotated training data in order to develop novel means of evaluation and comparison of trained models, and to exploit and integrate information from several modalities, i.e. different experimental measurements of the same sample. In particular, for the latter novel network architectures using multiple input sources will be investigated.

2.2.4 Registration

Co-localization and image registration play an important role in analyzing data acquired at different time points or with different modalities. Multimodal registration, i.e. the alignment of images from different scanner types, is in particular challenging because there is no direct way to compare the appearance to optimize a nonlinear transformation. In addition, the process is sensitive to large initial misalignments and can be time-consuming, especially for large and high-resolution images.

Recently, several deep learning methods have been proposed that jointly learn suitable image features and a nonlinear regression function to directly predict a displacement field given two images. They aim to replace the commonly used iterative registration process with learned convolution filters that are optimized using spatial transformer networks (differentiable image sampling) with respect to a supervisory task, which can be either based on ground truth correspondences, weaker structural label information or intrinsic mathematically derived metrics. Currently, most approaches are not well suited for large alignment [9] and do not offer a direct possibility to include network parts (architectural and pre-trained weights) from complementary tasks (e.g. segmentation).

In this project, we aim to first develop deep learning based registration approaches that enable semantic guidance from segmentation networks and landmark localization tasks and thereby learn modality-invariant and thus better comparable representations. For both auxiliary tasks, expert supervision can be more easily generated than (dense) correspondences. Furthermore, by using guidance from anatomical structures a canonical coordinate frame (e.g. defined by deep multimodal auto-encoders [10]) will be learned that facilitates the registration of scans acquired with large initial misalignments, e.g. due to very different fields of view (3D synchrotron tomography to 2D histological slices). Second, we will develop new deep learning architectures that separate the task of feature learning from the spatial alignment as

initially outlined in [11]). This will enable the incorporation and transfer of knowledge from trained networks for additional tasks, e.g. segmentation and image reconstruction or enhancement.

2.2.5 Multi-task Learning

The focus of our research project is the integration of multiple complementary tasks for biomedical image analysis into a common network architecture, which leverages the joint use of underlying concepts such as anatomy and shape while decoupling the differences in input appearance and output domain. Recent work in computer vision and machine learning has already started to develop concepts that separate stylistic elements of images from its content (or shape) [12]. In the context of multimodal and multi-task learning, the joint analysis of sound and video has led to highly accurate classifiers with almost no expert supervision [13]. In computer vision it was shown that a variety of inverse image enhancement tasks can be solved in a superior manner using a multitask network [14].

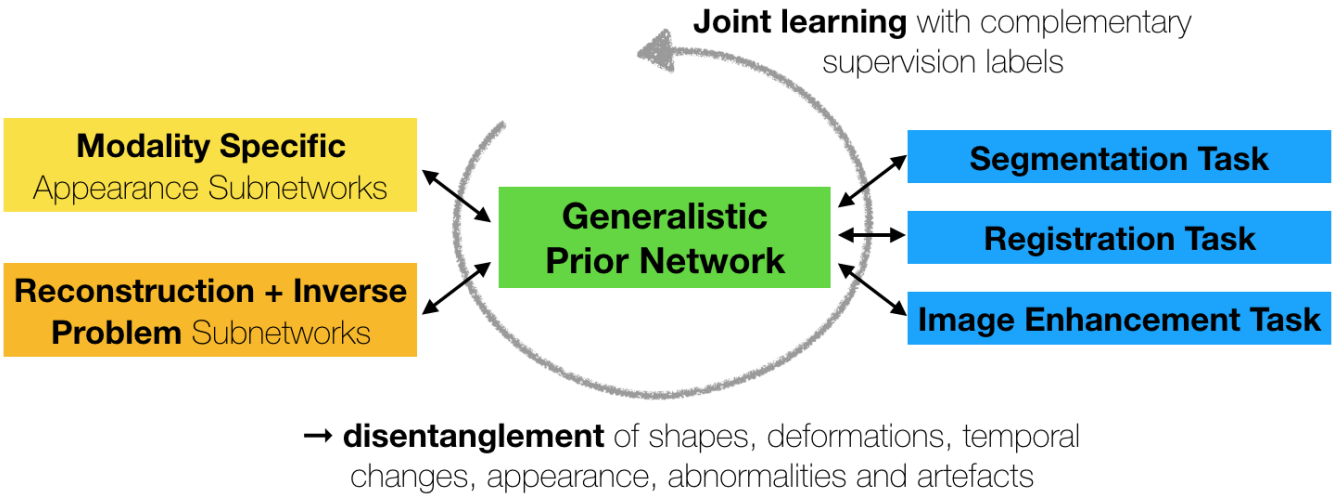

Figure 3: Proposed multi-task network model that benefits from shared priors common for several analyses and enables the disentanglement of task-specific differences for biomedical research objectives. Weights trained for the generalistic prior network are learned jointly from all complementary tasks to increase their ability to be transferred to new domains. The modelling of temporal changes, abnormalities or artifacts becomes easier and improves the image analysis.

Common in nearly all image analysis tasks is an implicit understanding of a deep network of a meaningful prior of the shape and content within all given datasets (e.g. from different modalities, but the same anatomical site). In this work package, we aim to develop a generalistic network architecture for biomedical images that learns a common shape and content prior based on multiple complementary tasks. In our prior work, an anatomically constrained network using a shape prior was successfully employed to improve motion artifact removal, image segmentation and super-resolution reconstruction [15]. Here, we aim to extend this concept to multimodal input as well as to images reconstructed with different levels of artifact correction (within inverse reconstruction tasks). On the other hand, we will implement several subtasks that will allow us to use complementary supervisory labels for image segmentation, registration and image enhancement. In general, we can observe that many work packages in this project will share similar CNN architectures (multiscale U-Net) or iterative refinement of the image analysis steps. These processing tasks are directly suitable for adapting state-of-the-art multi-task learning strategies, including cross-stitching networks, deep relationship networks or uncertainty-driven knowledge share, which are currently only employed in computer vision as well as attention mechanisms. This will enable us to jointly learn relevant image priors (of normal shapes) and disentangle them from modality specific appearance, artifacts and temporal changes of tissue and material composition to improve the accuracy and robustness of each individual task (see Figure 3).

2.2.6 Unified framework

The main intent of the unified framework is to enable scientist and users from the imaging community a quick and efficient application of deep learning methods for image processing and analysis in order to focus on scientific problems instead of implementation issues. While in the first step the focus is on image enhancement, segmentation and registration for specific problems, the idea of the framework is, to provide a modality and domain independent implementation that allows an easy integration of additional tasks and domain expert knowledge. Particular focus is on multi-task approaches in order to solve multiple processing steps simultaneously and/or in an iterative, alternating fashion. Therefore, we aim at a high-level formulation of the applied deep learning approaches such that method and application specific code has to written only once or can easily be adapted for other domains and modalities. Common python-based machine learning libraries such as TensorFlow or PyTorch will be used to facilitate fast prototyping and code development.

3 Financial framework; reasons for the necessity of the BMBF-funding

In total the consortiums applies for 1,170,157 EUR over the funding period of 3 years. Detailed justification of the funding requested is given in the accompanying data sheet. A request for funding of this project has not been submitted to any other addressee. In the case we submit such a request, we will inform the Federal Ministry of Education and Research immediately. There is no funding possible for our entirely Germany-based consortium by the European Commission such as by the Horizon 2020 program for Research and Innovation.

HZG: The division *Metallic Biomaterials* of the *Institute of Materials Research* belongs to the Helmholtz Program *Advanced Engineering Materials* within the Research Field *Key Technologies*. The corresponding funding is foreseen for materials development and biological assessment. The imaging activities of HZG relevant to this proposal are funded by two BMBF projects (05K16CGA & 05K16CGB). Available resources do not cover development and upgrade of currently available segmentation, registration and artifact reduction techniques or software. Therefore, this project can only be realized with external funds. For this project, HZG cannot apply at other national sources (e.g. DFG) because HZG as a Helmholtz Center is not eligible for funding of 'Einzelantrag'. The application at European funding agencies is not possible because the consortium is too small.

UzL: Very complex research projects which are carried out at the UzL cannot usually be realized with its own budget. An ambitious goal, as it is described in the present project, has risks in its implementation, which are reflected in particular in the resolution and artifact reduction to be achieved. The considerable research aspect will be taken over by a PhD student who will work full-time on the project. The resulting financial risk cannot be covered solely by IMT and can only be compensated by funding. Without the requested grant, processing would not be justifiable, as it would be accompanied by a significant reduction in resources.

DESY: The IT division at DESY, as well as several other scientific groups at DESY, are employing deep learning approaches in various scientific areas (see https://indico.desy.de/indico/event/21640/ for an overview of activities), and substantial efforts go into provisioning of high performance GPU compute environments. The ambitious goals of this project both in segmentation and provisioning of sustainable services cannot be covered by available resources or currently running third-party projects. The application at European funding agencies is not possible, since none of the open or published calls align with the methodologies and goals of the proposed project.

4 Quality and relevance of solid data basis for use in the project in development and validation and the accessibility at project start

In vivo MRI and µCT data of the same implants is currently collected by HZG in cooperation with the Molecular Imaging North Competence Center (MOIN CC) (BMBF project MgBone 05K16CGB). 60 animals are imaged every 30 days over a period of 120 and 150 days, respectively. SRµCT data of non-

degradable and degradable bone implants at various healing stages is available in large quantities from research projects carried out by HZG. Tomograms from about 120 critical point dried samples and 72 fresh samples of in situ biomechanical tests on implants are available from the BMBF project Synchroload (05K16CGA). 3D volumes are segmented manually with currently about 20 segmentations with high accuracy of large-scale tomographic volumes (1500^3 voxels) available. 53 segmented tomograms are available from the MSCA ETN MagnIM (289163). 77 SAXS regions from 33 samples have been registered with histology and SRµCT. Two PhD student of the Marie Skłodowska-Curie Action (MSCA) European Training Network (ETN) MgSafe will continue the work on the segmentation of tomograms and registration of SRµCT, SAXS and histology. 20 data sets from nano X-ray diffraction experiments are available which allow estimating phase compositions. Further information is available for some samples on the chemical composition from EDX and XRF and on the bone mineralization in 3D from SAXS tensor tomography. The data is stored on the DESY computer center and can be easily accessed via file transfer protocol. All project collaborators will be added to the pool of users for the respective data. HZG is currently developing and implementing a database containing (or linking to) experimental and simulation data from various modalities and sources. The database will store data in a way that allows easy access from different simulation and modelling tools as well as data processing and analysis tools. The database will contain data from MRI, SRµCT, labCT, optical imaging, SEM, TEM and other sources as well as materials data including microstructure and mechanical properties, processing data from different fabrications steps, in vitro degradation data, and biological and medical data.

Due to their strict post market surveillance, Syntellix actively collects and analyses patient data resulting from treatments with their products. Within a Syntellix initiated post market clinical follow up study, 90 patients undergo a scaphoid fracture treating surgery in which MAGNEZIX®-based compression screws are implanted. As part of the follow-up procedure of this study, clinical MRI data is achieved 6 weeks and 3, 6 and 12 months post-surgery. MRI data for further analysis will be provided in the clinically format.

In cooperation with the Department of Radiotherapy and the Clinic for Radiology and Nuclear Medicine at the University Medical Center S-H (UKSH) / Campus Lübeck, CT data is collected from patients with various metal implants. The constantly increasing number of data sets is anonymized and available as raw data and reconstructed data. Furthermore, it is possible to acquire specific data sets using multiple clinical CT devices (Biograph mCT and Somatom Definition AS, Siemens AG, Erlangen, Germany). Access to the raw data of the CT device often poses a major barrier, as these are usually stored in encrypted form and can only be read with special software. The applicant has signed a Non-Disclosure Agreement (NDA) with Siemens Healthcare GmbH since 2006. This refers in particular to readout routines provided by Siemens which allow accessing the raw data of the various Siemens CTs.

In the human medical domain, there has recently been a substantial growth in publicly available imaging datasets (CT, MRI, histology and more) that is provided with expert annotations. Exemplary are the medical decathlon datasets (http://medicaldecathlon.com) that contain 1746 usually high-resolution 3D scans acquired with CT or MRI of human anatomy and pathology spread across ten anatomical sights. Each scan is manually labelled by experts and may serve as basis for training deep networks for medical image that capture anatomical priors and ease the learning of task specific models. Furthermore the cancer imaging archive (https://www.cancerimagingarchive.net) has made available hundreds of medical scans with expert annotations and/or medical diagnoses and the German National Cohort (NaKo) currently acquires over 20,000 MRI scans of the general population with additional demographic and epidemiological parameters for which UzL has submitted a data access grant together with MEVIS Fraunhofer Bremen.

5 Relevance of the method / algorithm / software tool in (further) development, description of its impact on specified life-science field of application.

The specific tasks addressed within this project namely image enhancement, data segmentation and multimodal registration are common and often mandatory preprocessing steps for further analysis and interpretation of the acquired data in medicine, life science and other fields. In osteology e.g. CT imaging

techniques are used to assess the bone density and architecture which requires prior segmentation in order to determine the bone mass density or the structure and texture of the trabecular network [16]. In clinical MRI and CT e.g. the presence of metallic implants causes severe artifacts which limit the diagnostic image evaluation and demand for image enhancement techniques to generate artifact-free images [17,18]. There is a high demand for such tools among users of synchrotron radiation facilities such as PETRA III across disciplines including life sciences. E.g. almost all users of the imaging end stations of P05 and P07 at PETRA III have to perform at least one, often more, of these tasks.

Registration and segmentation of multimodal and/or longitudinal data sets are two of the major tasks in medical imaging and other fields which are typically affiliated with each other [19,20]. Each single task is very time-consuming and often requires weeks and months to fulfill. The processing is often impeded by imaging artifacts, low contrast and noise requiring further time-consuming data pre-processing. A multi-task learning approach solving these tasks simultaneous is therefore of great general interest for a wide user community.

6 Structure of the project, management and coordination

The project is structured in seven scientific and one management work packages (WP). WP1 (lead: HZG) is dedicated to data pre-processing. WP2 (lead: UzL), WP3 (lead: DESY), WP4 (lead: HZG) implement (single-task) deep learning approaches for image enhancement, segmentation and registration, respectively. WP5 (lead: UzL) is dedicated to multi-task solutions combining WP2, WP3 and WP4. WP6 (lead: HZG) is dedicated to the development and implementation of a unified framework for multi-task learning. WP5 and WP6 require a close collaboration and a steady exchange of knowledge between partners. WP7 (lead: HZG) is concern management issues.

HZG acts as interface between the partners in the consortium and is responsible for the coordination of the project and the monitoring of its progress. At project start, a kick-off meeting will be organized by HZG in order to coordinate tasks and discuss pathways of communication, data and knowledge exchange between partners. Each partner will participate in a monthly jour fix to discuss the progress of each WP. In quarterly virtual meetings the WPs will present their progress to the partner institutes. Every half year, the consortium meets in person to give project updates and to discuss adjustments to the work plan if needed. In accordance with the personal meeting, yearly workshops will be organized to present the algorithms developed and disseminate knowledge.

7 State of science

Deep learning methods for inverse problems

Recently, the field of inverse problems and in particular CT reconstruction algorithms has benefited from the constantly evolving area of deep learning techniques. The use of CNNs has shown a general potential to reduce artifacts in the field of noise suppression and super-resolution of natural and medical images and more specifically in low-dose imaging [21,22,23]. However, the potential for improvement in artifact reduction appears to be even greater as the strong disruption affects the entire reconstructed image. For example, a prior image is particularly important to compensate for ill-posed inverse problems. In contrast to classical priors, which are mostly based on the simple assumption of piecewise constant image areas (sparsity in gradients or wavelet coefficients) and also permit implausible solutions, deep convolutional networks (DCNNs) can be used to learn significantly more precise assumptions from realistic images.

First approaches for image restoration with neural networks were presented in [23], where an iterative nonlinear diffusion method with trainable activation functions in the form of a recurrent convolutional network was developed, which provided advantages for denoising as well as for the reconstruction of super-resolved images. In [24] this method was transferred to classical fully-convolutional NN architectures and simultaneously trained for a number of different (artificially added) noise levels with an

L1 norm. The noise suppression and enhancement of image resolution was particularly improved by learning with residual links.

As of recently, a number of approaches have been presented which, in CT reconstruction, suppress the unavoidable noise with a few low-dose projections using learning methods (see [25,26,27]) and are mostly built with multiscale CNN and using residual connections. In [14,22,27] an iterative optimization is used and a fully connected auto-encoder is trained unsupervised for a thinly populated low-dimensional image description. Due to the high number of trainable weights [21] is limited to 2D. In [27], the iterative reconstruction algorithm for limited X-ray projections is represented by an unrolled recurrent network, which is a single iteration of the reconstruction as layers in CNN, but finally only results in a common multiscale U-Net architecture [28] with residual connection. While classical learning methods have already been successfully used for CT images with metal artifacts [29,30], we do not yet know of any approach using deep learning (e.g. CNNs) in combination with classical metal artifact reduction methods and different image modalities. Even though different approaches for a CT reconstruction in combination with deep learning methods have already been presented, these methods either focus on CNNs that try to project the solution onto a set of desired reconstruction images [31], an intelligent parameter tuning of the optimization [32], or a CNN-based regularization term [33]. Nonetheless, a direct deep learning based reconstruction algorithm still needs to be found.

Segmentation

The task of segmentation is typically defined as identifying the set of pixels or voxels which make up either the contour or the interior of the object(s) of interest. In Life Sciences and especially in synchrotron imaging with an abundance of different imaging techniques (conventional absorption contrast, phase contrast techniques (propagation-based, grating interferometry, Zernike, etc.), micro-/nano-tomography, full-field/scanning methods, etc.), segmentation is typically done manually or in a semi-automatic fashion since fully automatized methods often fail because of the heterogeneity and complexity of acquired data sets. Therefore, a trade-off between quality and quantity of a manual or semi-automatic segmentation is usually required due to the typically large of amount of data sets. In addition, the quality of a manual segmentation strongly depends on the domain expert knowledge as well as the experience of the person segmenting the data.

Machine learning techniques (MLTs) perform automated segmentation by determining the optimal decision boundary in the high-dimensional space of discriminant features. The features extracted from images by human researchers are called handcrafted features. On the other hand, modern deep learning algorithms (DLAs) learn features that optimally represent the data for the problem at hand. Indeed, DLAs consist in models or networks composed of many layers that transform input data (e.g. images) to outputs (e.g. disease present/absent) while learning increasingly higher level features [34]. CNNs containing many layers that transform their input with convolution filters of a small extent are the most successful type of models for image analysis to date. Image segmentation can be performed by letting CNNs classify each pixel in the image individually. While this approach is computationally redundant, fully convolutional networks (fCNN) can be applied to an entire input image or volume in a much more efficient way and without any resolution penalty. The most well-known fCNN architectures is U-net [28] that is characterized by the combination of an equal amount of upsampling and downsampling layers and that provides a segmentation map directly in one forward step, and is therefore suitable for end-to-end learning of tissue segmentations and tissue boundary discrimination tasks. Moreover, a full 3D segmentation can be achieved by feeding U-net with a series of 2D annotated slices from the same volume by combining U-nets trained on orthogonal sections of the same data or eventually by employing 3D U-nets. In several other domains CNNs have been successfully used, e.g. to produce higher-quality, synthetic brain MRI to improve extraction of quantitative information from the acquired images [35], to provide automatic segmentation in MRI for knee cartilage [36] and proximal femur [37], and for nonlinear decomposition of CT images for metal artifact reduction [38]. Even though various deep learning architectures for semantic segmentation have been investigated, U-Nets are still the most widely used deep learning networks in biomedical image analysis [39]. The major challenge in the field of biomedical

image segmentation usually lies in the availability of a sufficient number of high-quality, well annotated training data. This is often addressed by enriching the training data with heavy augmentation.

Registration with deep learning

Compared to other fields relatively little research has yet been undertaken in deep learning based image registration [34] and most of this research has been published in the last two years. These methods mostly aim to learn a function in form of a CNN that predicts a spatial deformation which warps a so-called moving image to a fixed image. Based on how networks are trained, one can categorize these approaches into fully-supervised [40], unsupervised [9,41] and weakly-supervised [42,43] methods. *Supervised* methods require dense ground-truth deformation fields for training [44]. These deformation fields have to be either generated by synthetic deformation models or produced by classic image registration methods, which limits their accuracy by the performance of existing algorithms or simulations.

In contrast, *unsupervised* methods do not require any ground-truth data. The learning process is driven by image similarity measures or more general by evaluating the cost function of classic variational image registration methods [45]. An important milestone for the development of these methods was the introduction of the spatial transformer networks [46] for differentiable warping of moving images during training. Weakly-supervised methods also do not rely on ground-truth deformation fields but training is still supervised with prior information. The labels of the moving image are transformed by the deformation field and compared within the loss function with the fixed labels. All anatomical labels are only required during training. Combining weakly-supervised and unsupervised loss terms [45] further improves the alignment quality.

Multi-task analysis / Unified framework

While multi-task learning has attracted very much research in the computer vision field in recent years [47]. There are various parts of deep neural network architectures that can be leveraged for multi-task learning, since visual image understanding forms the basis and a common prior for many lower level analysis tasks. Deep relationship and deep adaptation networks e.g. learn to transfer network weights across domains by aligning the joint distributions of multiple domain-specific layers [48]. Cross-stitch networks enhance the knowledge transfer by learning layer and abstraction-level specific interchange weights between two networks [4]. Furthermore, multi-task learning was realized by uncertainty weighting of sub-network losses for scene geometry and semantic segmentation [49] and automatic search strategies over various multi-task modules [6].

In the medical domain comprehensive multi-task deep learning approaches are still rare. Transfer learning for domain adaptation has been extensively studied for pathology detection and segmentation in [50]. MLTs have also been applied to perform single- and cross-modality image analysis e.g. to reduce gadolinium dose in contrast-enhanced MRI [51] and to synthetize CT from MRI [52,53]. A deep neural network is used to learn an automated transform for data reconstruction by manifold approximation for data acquired from various modalities in [54]. Joint learning of motion and segmentation was studied for cardiac MRI images in [55]. Ultrasound-guided vein thrombosis diagnosis that relies on both landmark localization and deformation analysis was addressed using a multi-task deep network in [56]. Another area of interest that is not fully exploited for multi-task learning mechanisms are attention gated networks that have been successfully applied to medical image analysis (segmentation and localization) in [5].

8 Own previous work

HZG: The division *Metallic Biomaterials* of HZG covers the developmental chain from alloy design, production and processing via in vitro degradation studies towards cell culture experiments and animal trials. The correlation between the material properties and the environmental influence on degradation on one hand, and the tissue specific stimulation of cellular responses on the other hand, are in the focus of research activities. HZG contributes a long-standing experience in tomographic imaging and

multimodal investigations of biodegradable implants. Collaborative projects using synchrotron µCT and nanoCT are being undertaken to investigate the corrosion of Mg implants in situ, in vivo and on explants of implanted screws. The performed investigations include in vitro degradation tests, in situ biomechanical tests on implants using SRµCT, 2D SAXS and 3D SAXS/WAXS tensor tomography to probe the mineralization of bone, energy dispersive X-ray spectroscopy (EDX) to probe the chemical composition, X-ray diffraction (XRD) probing the phase, residual stress and texture of materials, and 3DXRD (grain mapping) to probe position, orientation and internal strains of grains within a gauge volume.

(1) Moosmann, J., Ershov, A., Altapova, V., Baumbach, T., Prasad, M. S., LaBonne, C., Xiao X., Kashef, J. & Hofmann R. (2013) X-ray phase-contrast in vivo microtomography probes new aspects of Xenopus gastrulation. *Nature* 497, 374–377.

(2) Moosmann, J., Ershov, A., Weinhardt, V., Baumbach, T., Prasad, M. S., LaBonne, C., Xiao X., Kashef, J. & Hofmann R. (2014). Time-lapse X-ray phase-contrast microtomography for in vivo imaging and analysis of morphogenesis. *Nature Protocols* 9(2), 294-304.

(3) Galli, S., Hammel, J., Herzen, J., Damm, T., Jimbo, R., Beckmann, F. Willumeit-Römer & R. (2016). Evaluation of the degradation behavior of resorbable metal implants for in vivo osteosynthesis by synchrotron radiation based x-ray tomography and histology. *Proc. SPIE - In Developments in X-Ray Tomography X*, 9967:13.

(4) Zeller-Plumhoff, B., Helmholz H., Feyerabend, F., Dose, T., Wilde, F., Hipp, A., Beckmann, F., Willumeit-Römer, R. & Hammel, J. U., (2017) Quantitative characterization of degradation processes in situ by means of a bioreactor coupled flow chamber under physiological conditions using time-lapse SRµCT, *Materials and Corrosion* 69:298-306.

(5) Moosmann, J, Zeller-Plumhoff, B., Wieland, D.C.F, Galli, S., Krüger, D., Dose, T., Burmester. H, Wilde. F., Bech. M, Peruzzi. N., Wiese, B., Hipp, A., Beckmann, F., Hammel, J.U. & Willumeit-Römer, R. (2017) Biodegradable magnesium-based implants in bone studied by synchrotron radiation microtomography. *Proc. SPIE - Developments in X-Ray Tomography XI*, 10391:23.

UzL-IMI: The research group of Prof. Mattias Heinrich at the Institute of Medical Informatics (IMI), currently comprises six PhD students that actively advance the field of deep machine learning in medical imaging with specific focus on image segmentation and registration. Most importantly, the new state-of-the-art multimodal registration metric MIND (modality independent neighborhood descriptor) was proposed in 2012 and a novel deformable convolutional architecture for scale- and rotation-invariant segmentation won the best paper award for Medical Imaging with Deep Learning in 2018, in addition many methodological advances in deep learning have been developed in collaboration with Imperial College London (Dr. Oktay).

(1) Heinrich, M. P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F. V., Brady, M. & Schnabel, J. A. (2012). MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical image analysis* 16(7), 1423-1435.

(2) Heinrich, M.P., Oktay, O. & Bouteldja, N. (2019). OBELISK-Net: Fewer Layers to Solve 3D Multi-Organ Segmentation with Sparse Deformable Convolutions. *Medical image analysis* 54(1)-9.

(3) Blendowski, M. & Heinrich, M. P. (2018). Combining MRF-based deformable registration and deep binary 3D-CNN descriptors for large lung motion estimation in COPD patients. *International journal of computer assisted radiology and surgery*, 1-10.

(4) Lucas, C., Kemmling, A., Bouteldja, N., Aulmann, L. F., Mamlouk, A. M. & Heinrich, M. P. (2018). Learning to predict ischemic stroke growth on acute CT perfusion data by interpolating low-dimensional shape representations. *Frontiers in neurology* 9.

(5) Oktay, O., Ferrante, E., Kamnitsas, K., Heinrich, M. & Kainz, B. (2018). Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. *IEEE transactions on medical imaging* 37(2), 384-395.

UzL-IMT: For more than a decade, the Institute of Medical Engineering (IMT) has been engaged in the reconstruction of images. In particular, in the field of imaging with the use of computed tomography,

there is a great level of expertise available. The most important work in recent years has been achieved in the fields of iterative reconstruction, metal artifact reduction, and noise suppression:

(1) Stille, M., Kleine, M., Haegele, J., Barkhausen, J. & Buzug, T. M. (2015). Augmented Likelihood Image Reconstruction*. IEEE Transactions on Medical Imaging* 35(1), 158–173.
(2) Levakhina, Y. M., Müller, J., Duschka, R. L., Vogt, F., Barkhausen, J. & Buzug, T. M. (2013). Weighted simultaneous algebraic reconstruction technique for tomosynthesis imaging of objects with high-attenuation features. *Medical Physics* 40(3), 031106.
(3) Kratz, B., Weyers, I. & Buzug, T. M. (2012). A fully 3D approach for metal artifact reduction in computed tomography. *Medical Physics* 39(11), 7042–7054.
(4) Weidinger, T., Buzug, T. M., Flohr, T., Kappler, S. & Stierstorfer, K. (2016). Polychromatic Iterative Statistical Material Image Reconstruction for Photon-Counting Computed Tomography. *International Journal of Biomedical Imaging*, 2016. Article ID 5871604, 15 pages.
(5) Heinrich, M. P., Stille, M., & Buzug T. M. (2018). Residual U-Net Convolutional Neural Network Architecture for Low-Dose CT Denoising. *Current Directions in Biomedical Engineering 4*(1), 297–300.

Furthermore, the project will make use of the following patent: Stille, M. & Buzug, T. M., “Method and apparatus for reducing artefacts in computed tomography images,” US20170150937A1, DE 10 2014 007 095.6, WO 2015 173 251 A1.

DESY: In the context of the Helmholtz Analytics Framework (HAF) project DESY IT has increased its activities in the development and implementation of deep learning algorithms. We have implemented, and trained a U-Net for the segmentation of 2D slices of SRμCT data. So far the availability of 3D information has only been used to the extent that training and predictions were run along the three axis of the volume and subsequently the three different predictions have been combined into one, considerably improving the accuracy of the segmentation.

Preliminary evaluations of predictions on test data show that the predicted segmentation is significantly improved i.e. more precise than the semi manual segmentation provided by domain experts, despite that the training data had numerous errors in the ground truth. Application of the trained model to the training data can provide a more accurate ground truth, which can be used in an iterative manner for retraining of the network, further improving the prediction quality.

Analysis of the probabilities with which each pixel was assigned to a label, their distribution, skewness and kurtosis suggest that these may be good indicators to assess the quality of the segmentation. The evaluation of these probabilities not only point to images where automatic prediction of the segmentation failed, but their pixelwise visualization even indicates areas within each image where the model fails.

Recent Publications:

(1) Kaira, C. S., Yang, X., De Andrade, V., De Carlo, F., Scullin, W., Gursoy, D. & Chawla, N. (2018) Automated correlative segmentation of large Transmission X-ray Microscopy (TXM) tomograms using deep learning*, Materials Characterization* 142, 203-210.
(2) Yang, X., De Andrade, V., Scullin, W., Dyer, E. L., Kasthuri, N., De Carlo, F. & Gursoy, D (2018) Low-dose x-ray tomography through a deep convolutional neural network, *Scientific reports* 8 (1), 2575.
(3) Yang, X., De Carlo, F., Phatak, C. & Gursoy, D. (2017) A convolutional neural network approach to calibrate rotation axis of x-ray computed tomography, *Journal of Synchrotron Radiation* 24, 469-475.

9 Project partners and their competence

In order to achieve the scientific aims of the project requires expert knowledge in deep learning, algorithm development, high performance computing, data acquisition and processing as well as input from domain experts. Each partner contributes with complementary competence within the consortium.

Mattias Heinrich (UzL-IMI) received his PhD from the University of Oxford in 2013 and has been actively involved for several years in the automatic interpretation and analysis of medical image data, image-guided medical diagnosis and treatment (e.g. CT stroke diagnosis and image-guided radiotherapy) and methods for deep machine learning, which provide a foundation for many tasks in this planned project. His research group consists of currently six PhD students, which are funded by two ongoing DFG projects HE 7364-1/2 and DFG HE 7364-2/1, one industrial PhD project and stipends, and actively develop new state-of-the-art methods for deep learning medical image analysis: segmentation, motion tracking [45,57] and interpretable visualization of prior shape models [10,58] and fusion of multimodal medical image registration. He has published in numerous high-ranking journals (19x with IF 3.5 or higher) and 14 papers at the premier international conference MICCAI *(Medical Image Computing and Computer Assisted Interventions)* with acceptance rates of only 30%. He has won a best paper award for deep learning in image registration [59] (senior author, BVM 2018, Erlangen), the prestigious MICCAI Young Scientist Award for medical multimodal image registration (first author, MICCAI 2011, Toronto) [60] with deep learning based extensions [11] and very recently the best paper award at MIDL 2018 *(Medical Imaging with Deep Learning)* for scale-adaptive deformable filter kernels [61] as first author with a prize of $9000. Mattias Heinrich's research also investigates new methods (partly in cooperation with Imperial College London) to improve neural networks for use with a limited number of annotated training data to increase the plausibility of the results by prior knowledge [15], visualize attention for interpretable results [5], by using limited computing capacities for large-scale deep learning inference [62] and for deep-learning based ultrasound-guided thrombosis diagnostics [56]. Mattias Heinrich has been selected to be part of the program committees for MICCAI in 2015 and 2018 and for MIDL in 2019.

Prof. Thorsten M. Buzug (UzL-IMT) received his PhD in 1993 in Applied Physics from the Christian-Albrechts-University of Kiel, Germany, where he worked in the field of signal processing applied to chaotic systems. He was the leader of the Philips Research Cluster Medical Image Processing and responsible for several projects in that field. In October 1998 he has been appointed as professor of Physics and Medical Engineering at the RheinAhrCampus Remagen. 2000-2004 he was head of the Academic Development Committee of the RheinAhrCampus Remagen. 2004-2006 he was head of the Joint Council of Campus Departments. In December 2006 he became Director of the Institute of Medical Engineering at UzL. He has published numerous journal articles, conference papers and books. He is member of the National Academy of Science and Engineering (acatech), German Physical Society (DPG), German Society of Biomedical Engineering (DGBMT), German Society of Nondestructive Testing (DGfZP), the IEEE and SPIE.

Helmholtz-Zentrum Geesthacht (HZG): The division Metallic Biomaterials led by Prof. Willumeit-Römer is focusing on the investigation of Mg-based alloys as temporary bone implant material. In particular, tailoring the degradation behavior of the alloy by adjusting the alloy composition, the processing route and the surface characteristics is a central research topic. The institute covers the full value chain from material design, production, processing, in vitro corrosion and cell culture up to in vivo studies in mice and rats (HZG outstations at the synchrotron PETRA III in Hamburg and the Molecular Imaging Center MOIN in Kiel). Thus, HZG provides extensive data sets including annotated data and domain expert knowledge. Due to the very close collaboration and joint projects with the division *Materials Physics*, which is operating instruments at large scale facilities for structural investigation, HZG has expertise in X-ray diffraction and imaging using synchrotron and laboratory sources including data acquisition and processing, segmentation, and registration.

Deutsches Elektronen Synchrotron DESY: DESY IT operates the high performance computing (HPC) resource for PETRA III, FLASH and the European XFEL with more than 22,000 CPU cores and 120 state of the art GPGPUs, presumably one of the largest photon science dedicated compute platforms in Europe (https://confluence.desy.de/display/IS/Maxwell). Access to HPC resources (with high-end visualization capabilities) will be provided to project partners as an in-kind contribution. Additional resources and access to models and data will also be provided in-kind through European Open Science Cloud (EOSC) portals and services. DESY IT has been contributing to a large of number of international

and national project in fields related to data lifecycle management, data preservation in particle physics, photon and neutron sciences or scientific computing. More recent activities involve the EOSC and Helmholtz projects like Amalea and the Helmholtz Analytics Framework (HAF) exploiting machine and deep learning in various fields from predictive maintenance to image processing. A recent workshop (https://indico.desy.de/indico/event/21640/) and machine learning seminar series (https://indico.desy.de/indico/category/641/) provide an overview of current activities and expertise.

10 Work plan and work packages

Overview of work packages (WP) with assignment of tasks (T) and deliverables (D):

WP1 Data pre-processing (lead: HZG)
Data from different modalities collected on different devices requires a common standard for data storage and access including metadata extraction (e.g. type of modality, resolution, etc.). All data sets will be integrated into a database which is currently developed by HZG.
T1.1: Harmonization of data sets with respect to data format, pixel/voxel size, normalization, etc.
T1.2: Database integration of data sets.
D1.1: Harmonized data sets.
D1.2: Data sets integrated into database.

WP2 Image enhancement (lead: UzL)
WP2 explores deep learning (DL) approaches for image enhancement, i.e. for the reduction of metal artifacts and for contrast enhancement of human and animal data exhibiting a low signal of biodegradable implants.
T2.1: CNN-based approaches for post-processing artifact reduction.
T2.2: Direct DL-based CT reconstruction.
T2.3: Iterative reconstruction algorithm for artifact reduction in combination with DL-based image and raw data processing.
T2.4: Quality assessment of reconstructed CT scans with adversarial networks.
T2.5: Integration of a non-local Prior Image-based regularization and implant-dependent weighting.
T2.6: Transfer learning between modalities and resolutions.
D2.1: Tool for image enhancement, based on CNN post-processing.
D2.2: Reconstruction algorithm for artifact reduction including noise and metal artifacts in ex and in vivo animal data from labCT.
D2.2: Enhanced spatial resolutions of in vivo labCT data.

WP3 Segmentation (lead: DESY)
WP3 explores DL-based approaches for segmentation of challenging data. This involves the segmentation of individual datasets as well as the development of multimodal segmentation networks using multiple data sources in parallel.
T3.1: Study of the robustness of CNNs towards weakly annotated ground truth data. We will use standard test-sets, e.g. MNIST, and introduce varying amounts of random and systematic false annotations to evaluate the performance of the standard examples in dependence on the fraction of false annotations.
T3.2: Development of semi-automatic tools based on CNNs with U-net topology to shorten the time required and to increase the accuracy of the manual segmentation of datasets.
T3.3: Iterative DL-based segmentation using CNN with U-net topology for automatic segmentation and analysis of high-resolution SRμCT data, using weakly annotated ground truth data. Different approaches to exploit the availability of 3D spatial information will be evaluated.
T3.4: DL-based segmentation using CNN with U-net topology for automatic segmentation and analysis of pre-clinical MRI.
T3.5: Quality criterion derived from probabilities of the trained network. The quality of manual annotations (segmentations) varies strongly in precision, spatial resolution and completeness because of the complexity, heterogeneity and huge extent of the data. A tool, based on the computed probabilities of

the network and their distribution, will be developed to automatically identify potential critical results in terms of entire 3D volumes, whole 2D images, as well as local patches within the images, which require a manual review.
D3.1: Segmentation tool.
D3.2: Segmentation of animal data from ex vivo SRμCT, SAXS and histology.
D3.3: Segmentation of animal data from in vivo MRI and in vivo CT.
D3.4: Segmentation of animal data from ex vivo SRμCT and in / ex vivo MRI / CT.
D3.5: Segmentation of longitudinal data from clinical MRI.

WP4 Registration (lead: UzL)
WP4 explores DL-based approaches for registration of multimodal and/or longitudinal data sets.
T4.1: Development of DL-based registration methods, which include semantic guidance from segmentation networks and landmark localization tasks.
T4.2: Learn modality-invariant features for SRμCT, SAXS, histology, in vivo MRI and CT.
T4.3: Employ shape encoder methods to define a canonical anatomical coordinate frame for large displacement alignment of medical scans including 2D to 3D.
D4.1: Registration tool for ex vivo and in vivo animal data from SRμCT, SAXS and histology.
D4.2: Performance evaluation for in and ex vivo animal data from MRI, CT and histology.
D4.3: Registration of longitudinal and multimodal clinical MRI as well as CT data to transfer network weights for multi-task strategy.

WP5 Multi-task learning (lead: UzL)
WP5 explores synergies by identification of generic structures (filters, topology, etc.) in the neural networks used in WP2-4 and aims to develop tools for simultaneous image enhancement (reduction of artifacts/noise, contrast enhancement, etc.), segmentation and registration using multi-task learning approaches.
T5.1: Identification of generic structures in the neural network from WP2-4.
T5.2: Enhanced segmentation and registration of ex vivo animal data via iterative multimodal segmentation.
T5.2: Development of a software tool for multi-task analysis for simultaneous image enhancement, segmentation and registration.
T5.3: DL-based multimodal segmentation with CNNs using registered (WP4) data derived from different experimental methods. We will investigate as well the use of the registered raw-data, i.e. the images, as well as using the individual segmentations derived from WP3 to obtain better segmentations.
D5.1: Software tool for multi-task analysis (enhancement, segmentation, registration).
D5.2: Multi-task analysis of ex vivo animal data from SRμCT, SAXS and histology.
D5.3: Multi-task analysis in vivo animal data from MRI and labCT.
D5.4: Multi-task analysis of animal data from ex vivo SRμCT and in / ex vivo MRI / labCT.
D5.5: Multi-task analysis of longitudinal data from clinical MRI.

WP6 Unified framework (lead: HZG)
The generic structures identified in WP5 will be analyzed and generalized in order to translate the methods and tools developed in WP2-5 to use cases of other modalities and domains.
T6.1: Development of a unified framework for multi-task analysis across modalities and domains.
T6.2: Deployment of the unified framework to the Maxwell cluster at DESY.
T6.3: Performance evaluation of the unified framework for domains and modalities other than in WP2-5.
D6.1: Unified framework.
D6.2: GitHub code repository.
D6.3: Performance evaluation.

WP7 Management (lead: HZG)
For the management of the project all partners provide full service of the financial and legal department. In addition the following issues will be accounted for:
T7.1: Organization of regular meetings and personal exchange between partners to ensure the most

efficient way of communication and work progress.
T7.2: Organization of three workshops during the project phase.
T7.3: Reporting to BMBF.
D7.1: Seven project meetings (kickoff meeting, six annual project meetings).
D7.2: In combination with the project meetings three workshops which will be organized.
D7.3: Scientific and financial reports.

	IMT	IMI	DESY	HZG	Sum (PM)
WP1 Data pre-processing	3	2	2	9	16
WP2 Image enhancement	17	2	1	2	22
WP3 Segmentation	3	2	10	2	17
WP4 Registration	1	10	1	2	14
WP5 Multi-task learning	4	12	2	8	26
WP6 Unified framework	2	2	8	12	24
WP7 Management	0	0	0	1	1
Sum (PM)	30	30	24	36	120

Table 1: Distribution of person month (PM) workload within work packages (WP).

11 Intended exploitation of the project results

All tools and methods developed within this project will be integrated in an open source software framework which will be publically available as a GitHub code repository. The framework will further be implemented on the Maxwell cluster at DESY and will thus be available to all users of PETRA III and XFEL. In general, the demand for such tools is very high at PETRA III and, in particular, for the user groups at the imaging end stations of P05 and P07 which have to deal with at least one, often several, of the tasks of segmentation, registration and image enhancement. In particular, the segmentation of huge amounts of 3D volumetric data (in the order of terabytes) is a very time-consuming process often delaying data analysis for weeks and months.

Within DESY there is a number of distinct experiments which will benefit from the developed and implemented methods. In particular in macromolecular structural biology the demand for reliable and accurate multimodal methods for 3D reconstruction and segmentation is huge. Data on macromolecular structures is measured on the campus with different methods ranging from synchrotron based X-ray crystallography to cryogenic electron microscopy, to SAXS and XFEL based methods. Joining the potential of all methods in an automatic manner to obtain a model of the molecules under investigation is so far not used routinely.

Methods developed for segmentation and image enhancement will be used by Syntellix in order to evaluate the performance of their magnesium-based alloy, MAGNEZIX®, in medical applications.

12 References

[1] R. Mittal, J. Morley, H. Dinopoulos, E. G. Drakoulakis, E. Vermani, and P. V. Giannoudis, “Use of bio-resorbable implants for stabilisation of distal radius fractures: the United Kingdom patients’ perspective.,” *Injury*, vol. 36, no. 2, pp. 333–338, Feb. 2005.

[2] S. Virtanen, “Biodegradable Mg and Mg alloys: Corrosion and biocompatibility,” *Mater. Sci. Eng. B*, vol. 176, no. 20, pp. 1600–1608, Dec. 2011.

[3] A. H. M. Sanchez, B. J. C. Luthringer, F. Feyerabend, and R. Willumeit, “Mg and Mg alloys: How comparable are in vitro and in vivo corrosion rates? A review,” *Acta Biomater.*, vol. 13, pp. 16–31, Feb. 2015.

[4] I. Misra, A. Shrivastava, A. Gupta, and M. Hebert, “Cross-Stitch Networks for Multi-task Learning,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 3994–4003.

[5] J. Schlemper, O. Oktay O, M. Schaap, M. Heinrich M, B. Kainz, B. Glocker, D. Rueckert, “Attention gated networks: Learning to leverage salient regions in medical images,” *Med. Image Anal.*, vol. 53, pp. 197–207, Apr. 2019.

[6] A. R. Zamir, A. Sax, W. Shen, L. J. Guibas, J. Malik, and S. Savarese, “Taskonomy: Disentangling Task Transfer Learning,” presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3712–3722.

[7] M. Stille, M. Kleine, J. Haegele, J. Barkhausen, and T. M. Buzug, “Augmented Likelihood Image Reconstruction,” *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 158–173, 2015.

[8] M. Stille and T. M. Buzug, “Augmented Likelihood Image Reconstruction with non-local prior image regularization,” in *Proc. 4th Intl. Mtg. on Image Formation in X-ray CT*, 2016, pp. 145–148.

[9] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Išgum, “A deep learning framework for unsupervised affine and deformable image registration,” *Med. Image Anal.*, vol. 52, pp. 128–143, Feb. 2019.

[10] N. Bouteldja, D. Merhof, J. Ehrhardt, and M. P. Heinrich, “Deep Multi-Modal Encoder-Decoder Networks for Shape Constrained Segmentation and Joint Representation Learning,” in *Bildverarbeitung für die Medizin 2019*, 2019, pp. 23–28.

[11] M. Blendowski and M. P. Heinrich, “Learning interpretable multi-modal features for alignment with supervised iterative descent,” Dec. 2018.

[12] A. Gilbert, J. Collomosse, H. Jin, and B. Price, “Disentangling Structure and Aesthetics for Style-Aware Image Completion,” presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1848–1856.

[13] R. Gao, R. Feris, and K. Grauman, “Learning to Separate Object Sounds by Watching Unlabeled Video,” presented at the Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 35–53.

[14] J. H. R. Chang, C.-L. Li, B. Poczos, B. V. K. V. Kumar, and A. C. Sankaranarayanan, “One Network to Solve Them All - Solving Linear Inverse Problems using Deep Projection Models,” *ArXiv170309912 Cs*, Mar. 2017.

[15] O. Oktay *et al.*, “Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation,” *IEEE Trans. Med. Imaging*, vol. 37, no. 2, pp. 384–395, Feb. 2018.

[16] H. K. Genant, K. Engelke, and S. Prevrhal, “Advanced CT bone imaging in osteoporosis,” *Rheumatol. Oxf. Engl.*, vol. 47, no. Suppl 4, pp. iv9–iv16, Jul. 2008.

[17] B. Hargreaves, P. W. Worters, K. B. Pauly, J. M. Pauly, K. M. Koch, and G. E. Gold, “Metal Induced Artifacts in MRI,” *AJR Am. J. Roentgenol.*, vol. 197, no. 3, pp. 547–555, Sep. 2011.

[18] M.-Y. Lee, K.-H. Song, J.-W. Lee, B.-Y. Choe, and T. S. Suh, “Metal artifacts with dental implants: Evaluation using a dedicated CT/MR oral phantom with registration of the CT and MR images,” *Sci. Rep.*, vol. 9, Jan. 2019.

[19] M. Erdt, S. Steger, and G. Sakas, “Regmentation: A New View of Image Segmentation and Registration,” *J. Radiat. Oncol. Inform.*, vol. 4, no. 1, pp. 1–23, Oct. 2017.

[20] D. Rueckert and J. A. Schnabel, “Registration and Segmentation in Medical Imaging,” in *Registration and Recognition in Images and Videos*, R. Cipolla, S. Battiato, and G. M. Farinella, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 137–156.
[21] D. Wu, K. Kim, G. E. Fakhri, and Q. Li, “Iterative Low-Dose CT Reconstruction With Priors Trained by Artificial Neural Network,” *IEEE Trans. Med. Imaging*, vol. 36, no. 12, pp. 2479–2486, Dec. 2017.
[22] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, “Generative Adversarial Networks for Noise Reduction in Low-Dose CT,” *IEEE Trans. Med. Imaging*, vol. 36, no. 12, pp. 2536–2545, Dec. 2017.
[23] Y. Chen and T. Pock, “Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1256–1272, Jun. 2017.
[24] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising,” *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[25] Q. Yang *et al.*, “Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1348–1357, Jun. 2018.
[26] K. Mentl *et al.*, “Noise reduction in low-dose CT using a 3D multiscale sparse denoising autoencoder,” in *2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*, Tokyo, 2017, pp. 1–6.
[27] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep Convolutional Neural Network for Inverse Problems in Imaging,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
[28] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241.
[29] Y. Huang, Y. Lu, O. Taubmann, G. Lauritsch, and A. Maier, “Traditional Machine Learning Techniques for Streak Artifact Reduction in Limited Angle Tomography,” in *Bildverarbeitung für die Medizin 2018*, 2018, pp. 222–227.
[30] Y. Zhang and H. Yu, “Convolutional Neural Network Based Metal Artifact Reduction in X-Ray Computed Tomography,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1370–1381, Jun. 2018.
[31] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, “CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018.
[32] C. Shen, Y. Gonzalez, L. Chen, S. B. Jiang, and X. Jia, “Intelligent Parameter Tuning in Optimization-Based Iterative CT Reconstruction via Deep Reinforcement Learning,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1430–1439, Jun. 2018.
[33] B. Chen, K. Xiang, Z. Gong, J. Wang, and S. Tan, “Statistical Iterative CBCT Reconstruction Based on Neural Network,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1511–1521, Jun. 2018.
[34] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
[35] Camilo Bermudez, Andrew J. Plassard, Larry T. Davis, Allen T. Newton, Susan M. Resnick, and Bennett A. Landman, “Learning implicit brain MRI manifolds with deep learning,” presented at the Proc. SPIE, 2018, vol. 10574.
[36] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, “Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, 2013, pp. 246–253.
[37] C. M. Deniz *et al.*, “Segmentation of the Proximal Femur from MR Images using Deep Convolutional Neural Networks,” *Sci. Rep.*, vol. 8, no. 1, p. 16485, Nov. 2018.
[38] H. S. Park, S. M. Lee, H. P. Kim, and J. K. Seo, “Machine-learning-based nonlinear decomposition of CT images for metal artifact reduction,” *ArXiv170800244 Phys.*, Aug. 2017.
[39] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image

Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015, pp. 234–241.
[40] M.-M. Rohé, M. Datar, T. Heimann, M. Sermesant, and X. Pennec, "SVF-Net: Learning Deformable Image Registration Using Shape Matching," in *Medical Image Computing and Computer Assisted Intervention − MICCAI 2017*, 2017, pp. 266–274.
[41] J. Krebs, T. Mansi, B. Mailhé, N. Ayache, and H. Delingette, "Unsupervised Probabilistic Deformation Modeling for Robust Diffeomorphic Registration," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018, pp. 101–109.
[42] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A Learning Framework for Deformable Medical Image Registration," *IEEE Trans. Med. Imaging*, pp. 1–1, 2019.
[43] Y. Hu *et al.*, "Weakly-supervised convolutional neural networks for multimodal image registration," *Med. Image Anal.*, vol. 49, pp. 1–13, Oct. 2018.
[44] T. Sentker, F. Madesta, and R. Werner, "GDL-FIRE 4D: Deep Learning-Based Fast 4D CT Image Registration," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 2018, pp. 765–773.
[45] A. Hering, S. Kuckertz, S. Heldmann, and M. P. Heinrich, "Enhancing Label-Driven Deep Deformable Image Registration with Local Distance Metrics for State-of-the-Art Cardiac Motion Tracking," in *Bildverarbeitung für die Medizin 2019*, 2019, pp. 309–314.
[46] M. Jaderberg, K. Simonyan, A. Zisserman, and Koray Kavukcuoglu, "Spatial Transformer Networks," in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 2017–2025.
[47] S. Ruder, "An Overview of Multi-Task Learning in Deep Neural Networks," *ArXiv170605098 Cs Stat*, Jun. 2017.
[48] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep Transfer Learning with Joint Adaptation Networks," *ArXiv160506636 Cs Stat*, May 2016.
[49] A. Kendall, Y. Gal, and R. Cipolla, "Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7482–7491.
[50] H. Shin *et al.*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016.
[51] E. Gong, J. M. Pauly, M. Wintermark, and G. Zaharchuk, "Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI," *J. Magn. Reson. Imaging*, vol. 48, no. 2, pp. 330–340, 2018.
[52] D. Andreasen, J. M. Edmund, V. Zografos, B. H. Menze, and K. V. Leemput, "Computed tomography synthesis from magnetic resonance images in the pelvis using multiple random forests and auto-context features," in *Medical Imaging 2016: Image Processing*, 2016, vol. 9784, p. 978417.
[53] T. Huynh *et al.*, "Estimating CT Image from MRI Data Using Structured Random Forest and Auto-context Model," *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 174–183, Jan. 2016.
[54] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018.
[55] C. Qin *et al.*, "Joint Learning of Motion Estimation and Segmentation for Cardiac MR Image Sequences," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 2018, pp. 472–480.
[56] R. Tanno *et al.*, "AutoDVT: Joint Real-Time Classification for Vein Compressibility Analysis in Deep Vein Thrombosis Ultrasound Diagnostics," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 2018, pp. 905–912.
[57] I. Y. Ha, M. Wilms, H. Handels, and M. P. Heinrich, "Model-Based Sparse-to-Dense Image Registration for Realtime Respiratory Motion Estimation in Image-Guided Interventions," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 2, pp. 302–310, Feb. 2019.
[58] C. Lucas, A. Kemmling, N. Bouteldja, L. F. Aulmann, A. Madany Mamlouk, and M. P. Heinrich,

"Learning to Predict Ischemic Stroke Growth on Acute CT Perfusion Data by Interpolating Low-Dimensional Shape Representations," *Front. Neurol.*, vol. 9, 2018.

[59] M. Blendowski and M. P. Heinrich, "Combining MRF-based deformable registration and deep binary 3D-CNN descriptors for large lung motion estimation in COPD patients," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 1, pp. 43–52, Jan. 2019.

[60] M. P. Heinrich *et al.*, "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Med. Image Anal.*, vol. 16, no. 7, pp. 1423–1435, Oct. 2012.

[61] M. P. Heinrich, O. Oktay, and N. Bouteldja, "OBELISK-Net: Fewer layers to solve 3D multi-organ segmentation with sparse deformable convolutions," *Med. Image Anal.*, vol. 54, pp. 1–9, May 2019.

[62] M. P. Heinrich, M. Blendowski, and O. Oktay, "TernaryNet: faster deep model inference without GPUs for medical 3D segmentation using sparse and binary convolutions," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 9, pp. 1311–1320, Sep. 2018.