Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Towards standardized quality control for AI systems in critical care

Type: Postdoc or PhD

Applicants

- Dr. Stefan Haufe (Charité, PTB, and TU Berlin)

Background

In intensive care units (ICUs), physiological parameters of patients are constantly monitored and combined with laboratory data and electronic health records to facilitate optimal and timely decision making, e.g., in cases of sudden deterioration. The wealth and inhomogeneity of the available data, however, makes it difficult for healthcare professionals to take all facets of a patient's condition into account. Artificial intelligence (AI) and machine learning (ML) based systems hold great promise to support clinical decision making in this context. Trained on historical data with known outcome, such systems can predict a patient's future clinical trajectory from the multivariate and multimodal data stream. Recently, retrospective studies have demonstrated the capability of ML models to predict the emergence of acute kidney injury (Tomašev et al., 2019), a variety of complications in cardiothoracic surgery (Meyer et al., 2018), as well as of lethal courses in Covid-19 pneumonia (Lichtner et al., 2020) with high accuracy. The prospective adoption of ML models in clinical practice, however, raises additional questions regarding the fairness, robustness, certainty, and comprehensibility of their decisions on a single case basis. Models should moreover be robust to moderate changes of the input data due to measurement errors, missing data, outliers as well as distributional shifts that can occur when transferring learned models to new clinical sites. It is also desirable that ML models provide well-calibrated estimates of the uncertainty of their individual predictions, as such information can be used to implement different (liberal vs. conservative) treatment strategies. A systematic protocol to benchmark the quality of ML models in critical care along these dimensions has not been provided yet.

Project Aim, Objectives and Program

The proposed project will carry out first steps towards a standardized quality assessment of machine learning approaches in critical care. To this end, the project will define a suite of benchmarks including quality criteria, reference problems, reference data, and reference implementations of ML models. The successful applicant will be supervised by Dr. Stefan Haufe (new professorship on machine learning and inverse problems at TU Berlin) and closely collaborate with the Institute of Medical Informatics at Charité. At a later stage, a larger community around the topic should be build (e.g. through public data analysis challenges). The project will primarily use large public data sets such as MIMIC III/IV (Johnson et al., 2016), the AmsterdamUMCdb, and HiRID (Hyland et al., 2020). Established and novel clinically relevant learning problems will be defined, such as the prediction of mortality, the prediction post-operative delir, and the prediction of individual treatment efficacies. Data from the different centers will be cleaned and harmonized. Reference prediction models from the literature will be implemented and their generalizability will be studied. Perturbed versions of the data will be

created to study the robustness of the predictions to outliers, missing data, noise, and distributional shifts.

Approaches to obtain uncertainty estimates for predictions will be implemented and validated. Recommendations on how to achieve balanced performance according to different criteria will be formulated. Eventually, novel prediction models based on state-of-the-art neural network architectures (e.g., recursive, invertible, self-calibrated, attention-based) should be developed and validated.

Available data

- MIMIC III (N=53,423 hospital admissions and ~60,000 ICU admissions at Beth Israel Deaconess Medical Center in Boston between 2001 and 2012, Boston), HIPAA compliant de-identification, data access through signed data use agreement.
- AmsterdamUMCdb (N=23,106 admissions of 20,109 patients admitted from 2003 to 2016 to the Amsterdam University Medical Center, GDPR compliant de-identification, access through signed user agreement.
- HiRID (N~33,000 ICU admission at Bern University Hospital between 2008 and 2016, time resolution of 2min). HIPAA and GDPR compliant de-identification, data access through signed data use agreement.
- Complementary data from Charité intensive care units

Collaboration

- Technische Universität Berlin, Fakultät IV
- Charité – Universitätsmedizin Berlin, Institute of Medical Informatics

Candidate Requirements

- MSc and/or PhD in computer science/mpathematics/statistics/physics
- Theoretical and practical expertise in machine learning/data science, experience with neural network architectures
- Fluency in SQL, Python, and at least one deep learning framework
- Practical computer/data science experience (e.g. shell scripting, git, docker, frontends)
- Ideally, experience with ICU/clinical routine data
- Enthusiasm, collaborative spirit, willingness to organize and contribute to community efforts to define and advance the field of AI in critical care.

References

[1] Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. Nature. 2019 Aug;572(7767):116-119.

[2] Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, Sündermann SH, Stamm C, Hofmann T, Falk V, Eickhoff C. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med. 2018 Dec;6(12):905-914.

[3] Lichtner G, Balzer F, Haufe S, Giesa N, Schiefenhövel F, Schmieding M, Jurth C, Kopp W, Akalin A, Schaller SJ, Weber-Carstens S, Spies C, von Dincklage F. Predicting lethal courses in critically ill COVID-19 patients using a machine learning model trained on patients with non-COVID-19 viral pneumonia. 2020. Submitted.

[4] Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. MIMIC-III, a freely accessible critical care database. Sci Data. 2016 May 24;3:160035.

[5] Hyland SL, Faltys M, Hüser M, Lyu X, Gumbsch T, et al. Early prediction of circulatory failure in the intensive care unit using machine learning. Nat Med. 2020 Mar;26(3):364-373.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Machine learning and uncertainty quantification for bioelectromagnetic inverse solutions and signal separation methods

Type: Postdoc or PhD

Principal Investigators

- Dr. Stefan Haufe (Charité, PTB, and TU Berlin)
- Dr. Tilmann Sander-Thömmes (PTB, AG 8.21)

Background

Magneto- and electroencephalography (M/EEG) are non-invasive techniques that sense brain activity from outside the head, providing insights into the brain's functioning in health and disease. From a clinical perspective, it is desirable to localize pathological activity to the generating brain structures, for example to inform surgical planning in drug-resistant epilepsy. For this, the physical process mapping neuronal currents to the M/EEG sensors needs to be inverted. Typical inversion schemes (e.g., Haufe et al., 2011) provide only point estimates or degenerate posterior distributions (e.g., Cai et al., 2020) that are unsuitable for quantifying the uncertainty of the estimates. A similar situation is encountered for multivariate statistical machine learning algorithms such as independent component analysis (ICA) that decompose the data into artefact- and brain-related components (e.g., Sander et al., 2010,. Delorme et al., 2012; Lueschow et al., 2015). ICA has been shown to work in practice even if the experimental data do not meet the statistical criteria (most importantly, number of sources and stationarity) upon which it is based. Such model violations might introduce bias or variance in the decomposition and a quantification of the uncertainty of ICA decompositions is needed. Currently, there exists no approaches to provide well-calibrated posteriors in the context of M/EEG source localization or statistical source separation. Neither exist approaches that would address the M/EEG inverse problem as a supervised prediction problem using synthetic ground-truth data as inputs for neural networks.

Project Aim, Objectives and Program

The first part of the project will develop and validate novel M/EEG inverse source imaging techniques with well-calibrated built-in uncertainty estimates. We will generate realistic synthetic EEG data to serve as a ground truth using an existing simulation framework (Haufe and Ewald, 2019). In addition, public simulated data will be used. We will develop novel hierarchical Bayesian techniques for proper uncertainty modeling as well as supervised neural network (e.g. graph NN) based approaches that are equipped with built-in uncertainty estimates (e.g., Lakshminarayanan et al., 2017, Ardizzone et al., 2019). The quality of these approaches compared to resampling approaches. To this end standardized performance metrics will be developed.

In the second part, source separation techniques will be benchmarked to identify the conditions in which ICA can reliably isolate the stationary brain sources. From the simulation results, we aim to infer the detectability of brain sources in existing MEG data with the option to extend these datasets for further verification. A flexible tool to assess the stability of ICA decompositions shall

be developed simulating cortical network structure with added white and non-white noise. The tool will incorporate classical methods such as resampling. Results from real data will be compared to the results obtained on simulated data using suitable metrics.

The project will leverage the expertise of the PIs and their extensive networks of collaborators. PTBs large database of MEG recordings is the result of two decades of neuroscience and clinical application research. Doctoral projects can be formally hosted at the novel professorship for machine learning and inverse problems (Dr. Haufe) at TU Berlin.

Available data

- RAMP challenge on source localization of MEG signals (N=50000, synthetic data),
- Berlin Brain Connectivity Benchmark,
- Several group studies SQUID-MEG, e.g., 16 subjects in "Cognitive processing during time reproduction", dataset with 125 channels and 1200 sec duration each, study with ethics and informed consent.
- Group study novel OPM on-scalp MEG with 13 subjects "Event related fields due to auditory stimuli", dataset with 30 channels, study with ethics and informed consent.

Collaboration

- Technische Universität Berlin
- BIFOLD Berlin
- Charité – Universitätsmedizin Berlin (Prof. Andrea Kühn, Prof. Surjo Soekadar)
- University of California San Francisco
- INRIA Paris
- Universität Ilmenau (Prof. Jens Haueisen)

Candidate Requirements

- MSc and/or PhD in computer science/mathematics/statistics/physics
- Mastery of the theory and practice of machine learning and multivariate statistics, experience with Bayesian modeling and contemporary (e.g., invertible, graph) neural network architectures
- Fluency in Matlab, Python, and at least one deep learning framework
- Ideally, experience with neuroscience, electrophysiological data
- Enthusiasm, collaborative spirit, genuine interest to advance the field

References

[1] Ardizzone L et al. . Analyzing inverse problems with invertible neural networks. In "Seventh International Conference on Learning Representations". 2019.

[2] Cai C, Hashemi A, Diwakar M, Haufe S, Sekihara K, Nagarajan SS. Robust estimation of noise for electromagnetic brain imaging with the champagne algorithm. Neuroimage. 2020 Oct 8;225:117411.

[3] Haufe S, Tomioka R, Dickhaus T, Sannelli C, Blankertz B, Nolte G, Müller KR. Large-scale EEG/MEG source localization with spatial flexibility. Neuroimage. 2011 Jan 15;54(2):851-9.

[4] Haufe S, Ewald A. A Simulation Framework for Benchmarking EEG-Based Brain Connectivity Estimation Methodologies. Brain Topogr. 2019 Jul;32(4):625-642.

[5] Lakshminarayanan B, Pritzel A, Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles. In "Advances in neural information processing systems". 2017; 6402-6413.

[6] Delorme A, Palmer J, Onton J, Oostenveld R, Makeig S. Independent EEG sources are dipolar. PLoS One. 2012;7(2):e30135.

[7] Lueschow A, Weber JE, Carbon CC, Deffke I, Sander T, et al. The 170ms Response to Faces as Measured by MEG (M170) Is Consistently Altered in Congenital Prosopagnosia. PLoS One. 2015 Sep 22;10(9):e0137624.

[8] Sander TH, Knösche TR, Schlögl A, Kohl F, Wolters CH, Haueisen J, Trahms L. Recent advances in modeling and analysis of bioelectric and biomagnetic sources. Biomed Tech (Berl). 2010 Apr;55(2):65-76.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Advancing the theory and practice of machine learning model explanations in biomedicine

Type: Postdoc or PhD

Applicants

- Dr. Stefan Haufe (Charité, PTB, and TU Berlin)

Background

In healthcare, as in other safety-critical domains, there is a strong desire – driven by clinical and scientific but also ethical and legal considerations – to understand how a given machine learning model arrives at a certain prediction for a particular data instance (e.g., a new patient), which motivates the field of explainable or interpretable artificial intelligence (xAI, e.g., Montavon et al., 2018). This problem is inherently unsupervised, which means that the ground-truth cannot, even retrospectively, be obtained in practice. Validation of unsupervised methods is a prerequisite for applying such methods in clinical contexts. This principle must also hold for xAI methods. However, due to the lack of ground-truth information in real data, the vast literature on xAI resorts to subjective qualitative assessments or surrogate metrics, such as relative prediction accuracy, to demonstrate the "plausibility" of the provided explanations. Moreover, there is no universally accepted definition of the "importance" of a feature that could be utilized to construct synthetic ground-truth data. Features are often considered important if their omission leads to a degradation of prediction performance. However, it has been pointed out that this definition is flawed, as it applies to noise features lacking any statistical relation to the prediction target (Haufe et al., 2014). While we have provided a simple remedy for linear learning problems, this problem is expected to be aggravated in use cases requiring non-linear prediction models, such as the classification or segmentation of radiological images using convolutive neural networks, for which no comparable remedy exists yet. Novel, theoretically founded, definitions of explainability along with appropriately designed synthetic ground-truth data are needed in order to benchmark existing xAI approaches as well as to drive the development of improved methods.

Project Aim, Objectives and Program

This project aims to advance both the theoretical foundation of xAI and the practical, in particular, clinical, utility of explanation methods. As such, it will extend prior and ongoing work in the group of Dr. Haufe at Charité Berlin (novel appointment as professor for machine learning and inverse problems at TU Berlin ongoing), and benefit from close interactions with domain experts at TU Berlin and the BIFOLD research center. We will develop novel, useful, definitions of feature importance that can be leveraged to generate synthetic ground-truth data. These data will be used to quantitatively assess the "explanation performance" of existing xAI methods such as layerwise relevance propagation (Bach et al., 2015), local surrogates (Ribeiro et al., 2016), PatternNet (Kindermans et al., 2017), SHAP scores (Lundberg and Lee, 2017). To this end, novel performance metrics will be developed.

We will create a benchmark suite of non-linear prediction problems where the set of important features is known a-priori. These problems will range from simple toy examples involving few variables to realistic settings mimicking clinical use cases such as image classification or segmentation tasks. Based on the results of our objective quantitative assessments, improved explanation methods should be developed for particular tasks as well as classes of machine learning approaches. The developed reference data and tools should be disseminated and further developed (e.g. by organizing symposia, public data analysis challenges) in a community-driven effort.

Available data

- Synthetic data (unlimited)
- T1-weighted MR images as basis for novel synthetic datasets, e.g. UK Biobank (N=1451)., Human Connectome Project (N>4000). All data are de-identified and can be accessed after signing a data use agreement.

Collaboration

- Machine Learning Group, Technische Universität Berlin
- Berlin Institute for the Foundations of Learning and Data (BIFOLD)
- Berlin Center for Advanced Neuroimaging (BCAN), Charité - Universitätsmedizin Berlin

Candidate Requirements

- MSc and/or PhD in mathematics/statistics/computer science/physics
- Strong expertise in the theory of machine learning and statistics
- Fluency in Python, experience with deep learning frameworks
- Good knowledge concepts of machine learning model interpretation and causal inference
- Solid computer/data science skills (e.g. shell scripting, git, docker, frontends)
- Critical, scientific mindset

References

[1] Bach S, Binder A, Montavon G, Klauschen F, Müller KR, Samek W. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. PLoS One. 2015 Jul 10;10(7):e0130140.

[2] Haufe S, Meinecke F, Görgen K, Dähne S, Haynes JD, Blankertz B, Bießmann F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage. 2014 Feb 15;87:96-110.

[3] Kindermans PJ, Schütt KT, Alber M, Müller KR, Erhan D, Kim B, Dähne S. (2017). Learning how to explain neural networks: PatternNet and PatternAttribution. International Conference on Learning Representations. 2018; arXiv preprint arXiv:1705.05598.

[4] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In Advances in neural information processing systems. 2017.

[5] Montavon G, Samek W, Müller KR. Methods for interpreting and understanding deep neural networks. Digital Signal Processing. 2018; 73:1-15.

[6] Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Invertible neural networks for resolving the hemodynamic inverse problem

Type: Postdoc or PhD

Principal Investigators

- Dr. Sebastian Heidenreich (PTB, AG 8.41),
- Prof. Gabriele Steidl (TU-Berlin)

Background

In medical diagnostics, where the health status of a patient is often determined with little or no intervention, measurands often obtained indirectly by solving a statistical inverse problem. Since variability in biological systems is typically large, the inverse problem is often ill-posed and uncertainty estimations are difficult. In recent years, efficient methods for uncertainty quantification for indirect measurements have been developed and applied [1]. To improve efficiency and computational speed, invertible neural networks (INNs) are successfully applied to estimate posterior distributions (distributions of the measurand) [2,3].

However, the state-of-the-art INNs do not account for measurement error models and estimates of significant hyperparameters, which leads to a systematic underestimation of obtained uncertainties (peak and multimodal distributions). First mathematical foundations for the treatment of multimodal posterior distributions were developed by Hagemann et al. [4].

In the proposed project, a general framework to determine uncertainties for Bayesian inversion for computational expensive systems using INNs will be developed. This includes the treatment of measurement errors, model errors and uncertainties caused by stochastic INNs.
The developed framework will be applied to hemodynamics (blood flow through the human body) to solve the statistical inverse problem from simulated and clinical data to provide more clinically relevant information for physicians from measurements.

Project Aim, Objectives and Program

The overall objective of the project is to develop a stochastic INN approach for dealing with measurement errors, model errors, hyperparameters, multimodal posterior distributions, and apply it to hemodynamics to obtain information of the cardiovascular system form blood pressure measurements for improved diagnosis.

The specific objects of the project are:

- Develop a stochastic INN approach for Bayesian inversion incorporating measurement and model errors.
- Develop tools to capture multimodal posterior distributions in the framework of stochastic INNs.
- Train stochastic INNs with virtual hemodynamics reference data and estimate physiological parameters and associated uncertainties from simulated measurement data.
- Determine distributions of cardiac and arterial parameters like elastic PWV, muscular PWV, elastic diameter, muscular diameter, heart rate, stroke volume and pheripheral resistance for different cohorts from real data.

Work program is divided into 5 work packages (WP).

WP1: Extending the state-of-the-art INN approach to solve high dimensional statistical inverse problems by including measurement errors, model errors and hyper-parameter estimations.

WP2: Use appropriate Lipschitz constants to learn transport maps from multimodal Gaussian distributions to multimodal posterior distributions.

WP3: Train the developed stochastic INNs through databases consisting of simulated data for estimates of cardiac and arterial parameters. At least two databases of virtually adults are available to train the extended stochastic INNs. The databases contain cardiac and arterial parameters as well as pulse velocities, pressure curves and PPG [5,6].

WP4: Validation of the INN with simulated reference data.

WP5: Determine cardiac and arterial parameters and associated uncertainties for different cohorts from blood pressure measurements using pre-trained stochastic INNs near real time. Uncertainties include model errors, measurement errors and INNs uncertainties.

For PhD students, fewer use cases are considered in WP3 and WP5. For postdocs in addition to WP1/2: Investigate mathematical properties of transport maps (e.g. push forward) and use these additional structures to efficiently train the INN.

Available data

- Simulated data (open access): King's College London [5].
- Simulated data (open access): DOI: 10.5281/zenodo.3275625 [6].
- Real data (open access): MIMIC-III Waveform Database Matched Subset,

Collaboration

- TU-Berlin, WIAS, ZIB

Candidate Requirements

- PhD in applied mathematics, computer science, or similar
- Experience in at least one of the fields of deep learning or simulation
- Software experience: Python (preferably PyTorch) and joint development (Git)

References

[1] Heidenreich S., Gross H. and Bär M., *Bayesian approach to determine critical dimensions from scatterometric measurements,* Metrologia 55 S201 (2018).

[2] Ardizzone L., Kruse J., Rother C., and Köthe U., *Analyzing inverse problems with invertible neural networks.* In 7th International Conference on Learning Representations, New Orleans, Conference Track Proceedings. ICLR, 2019; Kruse J., Detommaso G., Schleichl R., and Köthe U., Hint: *Hierarchial invertible neuraltransport for density estimation and bayesian infernce* (2020), arXiv:1905.10687.

[3] P. Hagemann, N. Farchmin, A. Anderle, V. Soltwisch, S. Heidenreich, G. Steidl, *Invertible Neural Networks versus MCMC for Posterior Reconstruction in Grazing Incidence X-Ray Fluorescence*, in prep.

[4] Hagemann P. and Neumayer S. *Stabilizing invertible neural networks using mixture models* (2020), arXiv:2009.02994.

[5] Willemet M., Chowienczky P. and Alastruey J. *A database of virtual healthy subjects to assess the accuracy of foot-to-foot pulse wave velocities for estimation of aortic stiffness*, AJP Hear. Circ. Physiol. 309, H663-H675, 2015.

[6] Charlton P. H., Mariscal Harana, J., Vennin, S., Li, Y., Chowienczyk, P. & Alastruey, J., *Modelling arterial pulse waves in healthy ageing: a database for in silico evaluation of haemodynamics and pulse wave indices,* AJP Hear. Circ. Physiol., 317(5), pp.H1062-H1085, 2019.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Robust ML-based quantitative MRI

Type: Postdoc or PhD

Principal Investigators

- Christoph Kolbitsch (PTB, AG 8.13)
- Andreas Kofler (PTB, AG 8.13)

Background

Quantitative MRI provides a wide range of different physical and biophysical parameter such as relaxation times (T1, T2) or blood perfusion measurements[1]–[3]. Commonly, quantitative parameters are obtained from reconstructed images. Model-based image reconstruction allows to directly reconstruct quantitative parameters from the raw MR data and enables accurate quantification even from highly undersampled data[4]–[6]. Nevertheless, this approach has so far not found widespread use due to long reconstruction times.

Recently, ML has also been proposed for quantitative MRI. One of the main challenges of ML for quantitative MRI is the lack of ground truth experimental data. In addition to the common confounding factors such as physiological motion, scan duration of quantitative MRI can be further limited by contrast dynamics. Therefore, fully sampled reference data cannot be acquired.

In quantitative MRI, ML is commonly used instead of the data fit to allow for fast parameter estimation[7]–[10]. Although this leads to a speed-up of the parameter estimation, it does not ensure any data consistency with the raw data. This can be problematic if simulated data has to be used for training. Any differences between the signal model used for data simulation and the actual data acquisition can lead to unpredictable results. Several ML approaches have been presented which add a physics-based model to the training to overcome this challenge[11], [12]. Nevertheless, this can be computationally very demanding and hence is often only applied for proof-of-principle examples.

Most ML-based approaches for quantitative MRI only yield a point estimate and without information about the uncertainty or reliability of the obtained parameters. Especially for applications in the heart or liver this is important, because the accuracy of parameter estimation can highly differ between patients due to strongly varying anatomy or uncompensated physiological motion.

Project Aim, Objectives and Program

In this project, we want to develop a robust physics-guided ML approach for MR parameter estimation for clinically relevant 2D and 3D non-Cartesian data acquisition by completing the following tasks:

- Based on our previous work on ML-based image reconstruction for dynamic cardiac cine MRI[13], [14], we will combine MR image reconstruction and parameter estimation as a task-based ML approach with a physics-based data consistency term.
- Explore different approaches for transfer-/refinement-learning in order to overcome the challenge of limited training data by pre-training the network with simulated data.
- Extend the ML approach to allow for the estimation of a measure of uncertainty.

WP1: Tasked-based ML combines two or more networks such that the output of the intermediate networks is the optimal input of the subsequent network[12]. In this case here, we will combine a reconstruction network for 2D non-Cartesian scans with a network for parameter estimation. This joint end-to-end training ensures images which provide the best image features for parameter estimation.

WP2: Initial training will be carried out based on simulated data. Different strategies to adapt this pre-trained model for experimental data will be compared.

WP3: The developed model will be evaluated for 2D cardiac T1 mapping and MRF parameter estimation.

WP4: The approach will be extended from 2D non-Cartesian to a 3D non-Cartesian data acquisition to ensure these large datasets can be handled on a GPU and that training is still carried out in an efficient way.

WP5: In addition to providing quantitative parameters, the developed ML approach will be extended to also yield a measure for uncertainty.

WP6: The uncertainty estimate will be utilised to detect any outliers in the data (e.g. due to incomplete respiratory motion compensation).

The work packages defined above are aimed for a post doc. For a PhD student, the focus will be on WP1, 2, 3 and 5.

Available data

- Numerical simulation environment for T1, MRF and DCE-MRI[15]
- 2D Cardiac T1 mapping (20 scans in volunteers (different orientations), 50 patient scans)[16], [17]
- 2D Cardiac MRF (10 volunteers)
- 3D DCE-MRI of the liver (10 patients, 30 pre-clinical data sets (multiple time points))[18]

Collaboration

- King's College London (Claudia Prieto, MR image reconstruction using ML)
- University of Innsbruck (Markus Haltmeier, ML for inverse problems)
- Technical University of Munich (Markus Makowski, radiologist)
- Charité – Universitätsmedizin Berlin (Jeanette Schulz-Menger, cardiologist)

Candidate Requirements

- MSc in physics, electrical engineering, applied mathematics or in a comparable field of engineering/natural sciences
- Experience in one or more of the fields: signal and image processing, ML for inverse problems

References

[1] M. Jerosch-Herold, N. Wilke, A. E. Stillman, and R. F. Wilson, "Magnetic resonance quantification of the myocardial perfusion reserve with a Fermi function model for constrained deconvolution," *Med. Phys.*, vol. 25, no. 1, pp. 73–84, Jan. 1998.

[2] D. Ma *et al.*, "Magnetic resonance fingerprinting.," *Nature*, vol. 495, no. 7440, pp. 187–192, 2013.

[3] D. R. Messroghli, A. Radjenovic, S. Kozerke, D. M. Higgins, M. U. Sivananthan, and J. P. Ridgway, "Modified look-locker inversion recovery (MOLLI) for high-resolution T 1 mapping of the heart," *Magn. Reson. Med.*, vol. 52, no. 1, pp. 141–146, 2004.

[4] V. T. Olafsson, D. C. Noll, and J. A. Fessler, "Fast Joint Reconstruction of Dynamic R_2^* and Field Maps in Functional MRI," *IEEE Trans. Med. Imaging*, vol. 27, no. 9, pp. 1177–1188, Sep. 2008.

[5] K. T. Block, M. Uecker, and J. Frahm, "Model-Based Iterative Reconstruction for Radial Fast Spin-Echo MRI," *IEEE Trans Med Imaging*, vol. 28, no. 11, pp. 1759–1769, 2009.

[6] O. Maier *et al.*, "Rapid T 1 quantification from high resolution 3D data with model-based reconstruction," *Magn. Reson. Med.*, vol. 81, no. 3, pp. 2072–2089, 2019.

[7] J. Shao, V. Ghodrati, K. Nguyen, and P. Hu, " Fast and accurate calculation of myocardial T 1 and T 2 values using deep learning Bloch equation simulations (DeepBLESS) ," *Magn. Reson. Med.*, no. June 2019, pp. 1–15, 2020.

[8] Z. Fang *et al.*, "Deep Learning for Fast and Spatially Constrained Tissue Quantification from Highly Accelerated Data in Magnetic Resonance Fingerprinting," *IEEE Trans. Med. Imaging*, vol. 38, no. 10, pp. 2364–2374, 2019.

[9] O. Cohen, B. Zhu, and M. S. Rosen, "MR fingerprinting Deep RecOnstruction NEtwork (DRONE)," *Magn. Reson. Med.*, vol. 80, no. 3, pp. 885–894, 2018.

[10] Y. Bliesener, J. Acharya, and K. S. Nayak, "Efficient DCE-MRI parameter and uncertainty estimation using a Neural Network," *IEEE Trans. Med. Imaging*, vol. 39, no. 5, pp. 1–1, 2019.

[11] F. Liu, L. Feng, and R. Kijowski, "MANTIS: Model-Augmented Neural neTwork with Incoherent k-space Sampling for efficient MR parameter mapping," *Magn. Reson. Med.*, vol. 82, no. 1, pp. 174–188, 2019.

[12]H. Jeelani, Y. Yang, R. Zhou, C. M. Kramer, M. Salerno, and D. S. Weller, "A Myocardial T1-Mapping Framework With Recurrent and U-Net Convolutional Neural Networks," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1941–1944.
[13]A. Kofler, M. Dewey, T. Schaeffter, C. Wald, and C. Kolbitsch, "Spatio-Temporal Deep Learning-Based Undersampling Artefact Reduction for 2D Radial Cine MRI With Limited Training Data," *IEEE Trans. Med. Imaging*, vol. 39, no. 3, pp. 703–717, Mar. 2020.
[14]A. Kofler *et al.*, "Neural networks-based regularization for large-scale medical image reconstruction," *Phys. Med. Biol.*, vol. 65, no. 13, p. 135003, Jul. 2020.
[15]J. Mayer *et al.*, "Flexible numerical simulation framework for dynamic PET-MR data," *Phys. Med. Biol.*, vol. 65, no. 14, p. 145003, Jul. 2020.
[16]K. M. Becker, J. Schulz-Menger, T. Schaeffter, and C. Kolbitsch, "Simultaneous high-resolution cardiac T1 mapping and cine imaging using model-based iterative image reconstruction," *Magn. Reson. Med.*, vol. 81, no. 2, pp. 1080–1091, 2019.
[17]K. M. Becker *et al.*, "Fast myocardial T 1 mapping using cardiac motion correction," *Magn. Reson. Med.*, vol. 83, no. 2, pp. 438–451, Feb. 2020.
[18]M. Ippoliti *et al.*, "3D nonrigid motion correction for quantitative assessment of hepatic lesions in DCE-MRI," *Magn. Reson. Med.*, vol. 82, no. March, pp. 1–14, 2019.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Active learning for medical imaging using Fisher information

Type: Postdoc or PhD

Principal Investigators

- Jörg Martin (PTB, AG 8.42)
- Clemens Elster (PTB, AG 8.42)

Background

Active learning aims at selecting training data in a way such that the robustness of the trained net is maximized and/or the required number of training data is kept at a minimum. Active learning is a current line of research in deep learning (see [1-3]) and particularly relevant in cases where labeling is expensive or time consuming. This is particularly relevant in medical applications, where human experts are needed. State-of-the-art approaches are based on minimizing predictive uncertainty [2] or optimal exploration of feature space [3].

PTB-8.42 has implemented and tested current approaches of active learning for regression problems related to the contrast-detail curve determination in mammography image quality assessment. While clear benefits have been observed in some cases, the approaches often lack robustness, and in some cases even fail to improve training performance. Furthermore, the number of initial, randomly chosen, training samples has a large impact and strategies for its determination are still lacking.

Active learning is closely related to experimental design in statistics [4] for which the Fisher information matrix plays a key role. Current active learning approaches do not directly refer to those concepts, and the Fisher information matrix is rarely used in this context. One reason is that the calculation of the Fisher matrix is practically impossible for typical neural nets.

PTB-8.42 has successfully made use of the Fisher information matrix for detecting adversarial examples and out-of-distribution behavior [5,6]. The key idea was to use suitably selected projections of the Fisher matrix which scale well and can be efficiently calculated even for many millions of unknowns in the neural net.

Project Aim, Objectives and Program

The goal of the project is to design methods for active learning based on the Fisher information and to provide means for their efficient implementation. Specific objectives are to explore

- appropriate directions of projection of the Fisher information matrix for optimal active learning together with efficient calculation schemes,
- strategies for the choice of initial samples prior to starting the active learning strategy based on Fisher information,
- potential benefits of the Fisher information matrix approach by comparison with current active learning strategies for the two applications (i) mammography image quality assessment and (ii) ECG diagnostics.

Work program

- Preparation of databases for assessing active learning strategies for the two considered applications, and selection of network architectures to be considered.
- Development of suitable projections of the Fisher information matrix and efficient calculation schemes.
- Application of the developed active learning strategy based on the Fisher information for the two applications and comparison with results obtained by current approaches.

Available data

- Virtual mammography phantom data (>40 k, > 1 M possible), already available
- Labeled ECG data base published by PTB (> 20 k), publicly available

Candidate Requirements

- PhD and/or MSc in applied mathematics, computer science, or similar
- Experience in at least one of the fields of deep learning or statistics
- Software experience: Python (optimally in PyTorch) and joint development (Git)

References

[1] Mizokami S. *Deep active learning from the perspective of active learning theory*. Deep active learning. Springer, Singapore, 2018. 79-91.
[2] Gal Y, Islam R, Ghahramani Z. *Deep Bayesian Active Learning with Image Data*. ICML 2017: 1183-1192
[3] Wu D, Lin C-T, Huang J. *Active learning for regression using greedy sampling*. Inf. Sci. 474: 90-105 (2019)
[4] Pukelsheim F. *Optimal design of experiments*. Society for Industrial and Applied Mathematics, 2006.
[5] Martin J, Elster C. *Inspecting adversarial examples using the fisher information*. Neurocomputing 382: 80-86 (2020)
[6] Martin J, Elster C. *Detecting unusual input to neural networks*. Applied Intelligence (2020): 1-12.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Uncertainty in deep learning versus conventional statistics for applications in medicine

Type: Postdoc or PhD

Principal Investigators

- Jörg Martin (PTB, AG 8.42)
- Clemens Elster (PTB, AG 8.42)

Background

Reliable uncertainty quantification of deep learning approaches is particular important in medical applications. One source of uncertainty in a prediction of a trained net is the uncertainty of these parameters that remains after training. The huge number of parameters, however, prevents the application of established methods from conventional statistics such as the calculation of a covariance matrix or a full Bayesian inference.

Uncertainty quantification in deep learning is not based on a rigorous application of established methods from conventional statistics but either relies on some sort of approximation (e.g. variational Bayes [1], dropout [2]) or heuristic approaches such as ensemble learning [3]. The relationship of these approaches to methods from conventional statistics in situations where both approaches are applicable has not yet been fully explored nor is the validity of the approximation well understood.

One goal of this project is to explore this relationship for low-dimensional regression problems that can be treated using established methods from statistics as well as regularized deep learning approaches. Furthermore, research connecting deep learning with Gaussian process modeling [4] or non-parametric Bayes [5] shall be deepened. The overall goal is to be enhance the understanding of current uncertainty quantification in deep learning against the background of approaches applied in conventional (Bayesian) statistics. In addition to gaining fundamental insights in uncertainty quantification for deep learning, the results of this project will help in bridging classical uncertainty evaluation in metrology and in deep learning. Another goal is to apply the studied uncertainty methods to a regression problem in mammography image quality assessment.

The goal shall build on PTB-842's work done so far in implementing and testing current uncertainty quantification approaches for deep learning in regression problems and their application in case studies [6], along with PTB's broad experience in Bayesian inference and uncertainty evaluation in metrology [7,8].

Project Aim, Objectives and Program

The goal of the project is to deepen the understanding of the relationship between approaches currently applied for uncertainty quantification in deep learning and conventional Bayesian inference.

Specific objectives are

- Comparison of uncertainty characterization using conventional Bayesian statistics for low-dimensional regression problems with those obtained by current uncertainty approaches for deep learning.
- Deepen the understanding of uncertainty quantification in deep learning from the point of view of non-parametric Bayes and Gaussian process modeling.
- Provide guidance as to the choice of uncertainty quantification method in deep learning from the point of view of its consistency with current uncertainty evaluation standards in metrology.

Work program

- Selection of many (>5) regression scenarios for numerical comparisons including quality assessment for mammography; implementation of current state-of-the-art uncertainty quantification in deep learning; reference treatment of selected benchmark problems using Bayesian inference and different assumptions about available prior knowledge; comparison and assessment of results obtained by the statistical treatment and by deep learning.
- Adaption of deep learning approaches to account for prior knowledge such as physical constraints or prior knowledge about the regression curve. Application to benchmark problems and comparison with reference method (using conventional Bayes).
- Exploration of relationship between current approaches for uncertainty quantification in deep learning with non-parametric Bayes and deep Gaussian process modeling.

Available data

- Virtual mammography phantom data (>40 k, > 1 M possible), already available
- Simulated data for low-dimensional regression problems, readily available

Candidate Requirements

- PhD and/or MSc in statistics or mathematics
- Experience in statistics and, optimally, in deep learning
- Software experience: Python (optimally in PyTorch) and joint development (Git)

References

[1] Bishop, CM. *Pattern recognition and machine learning*. Springer 2006.
[2] Kingma, DP, Salimans T, Welling M. *Variational dropout and the local reparameterization trick*. Advances in neural information processing systems. 2015.
[3] Lakshminarayanan, Balaji, Pritzel A, Blundell C. *Simple and scalable predictive uncertainty estimation using deep ensembles*. Advances in neural information processing systems 30:6402-6413, 2017.
[4] Khan MEE. *Approximate inference turns deep networks into Gaussian processes*. Advances in neural information processing systems. 3091-3104, 2019.
[5] MacEachern, SN. *Nonparametric Bayesian methods: a gentle introduction and overview*. Communications for Statistical Applications and Methods 23 (6): 445-466, 2016.
[6] Hoffmann L, Fortmeier I, Elster C. *Ensemble learning for computational optical form measurement.* Submitted to *Sensor and Measurement Science International 2021*.
[7] Elster C et al. *A guide to Bayesian inference for regression problems*. European Metrology Research Program NEW04, http://www.ptb.de/emrp/new04.html 2015.
[8] Elster C. *Bayesian uncertainty analysis compared with the application of the GUM and its supplements*. Metrologia, 51(4), S159-S166, 2014.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Artificial intelligence and metabolite markers in the diagnosis and prognosis of Parkinson's disease

Type: Postdoc

Principal Investigators

- Prof. Dr. Gavin O'Connor (PTB, AG 3.24)
- Prof. Dr. Karsten Hiller (Technische Universität Braunschweig)
- Prof. Dr. Tim Kacprowski (Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School)

Background

Progressive neurodegenerative diseases, such as Parkinson's disease (PD) often present heterogenous symptoms making clinical diagnosis and patient prognosis difficult. The clinical diagnosis of PD relies mainly on the physical symptoms of the disease combined with medical history and a patient's response to dopamine treatment. However, the clinical manifestations are very heterogeneous and often lag far behind pathological changes, which can lead to a late diagnosis of the disease, which is known to impact on disease progression. The exact causes of PD are still providing a fertile area for clinical research and are thought to involve both genetic and environmental contributions, which somewhat explains the diversity of symptoms and differences in the rate of disease progression. This diversity leaves patients uncertain about their long-term quality of life and hampers objective assessment of clinical trials.

Metabolomics aims to detect global differences between samples based on the numerous metabolites, followed by datamining and bioinformatics. The metabolites detected are not only endogenous but also consist of co-metabolites from gut biota as well as exogenous sources, such as ingested pharmaceuticals and environmental chemicals. Current methods are capable of distinguishing between minor changes in endogenous or exogenous stimulus, enabling a link between genetic, environmental, and physical status, to specific pathological states. The last decade has seen advances in the biological sample treatment and instrumentation that now enable the identification and quantification of metabolites on a global scale. This combined with the increased robustness of analytical platforms and robotic based automation enables the high throughput analysis of a broad range of metabolites from a range of sample types.

Currently PTB provides reference measurements for a discrete set of metabolites based on the accurate quantification of a single metabolite (single targeted quantification method) the proposed approach uses multi-parameter analysis and the standardisation of these approaches is still required. This is the focus of a new Junior research group at TU Braunschweig and the impact and use of standardisation approaches in data collection will be incorporated in this project

Project Aim, Objectives and Program

The main task will be the use of AI approaches for the identification of metabolite signals that result in the early diagnosis, stratification, prognosis, and progression of PD patient cohorts. Subtasks will include: the identification of single and multiple metabolite markers that could

enable the early diagnosis of PD and stratification of differen phenotypes; the investigation of the synergistic effects of the use of multiple metabolite markers in predicting disease progression; an assessment of the impact of improved quantitative uncertainty estimation on AI protocols and their ability to stratify patient cohorts.

Working with Prof. Brit Mollenhauer, who initiated and heads the DeNoPa (*de novo* Parkinson) cohort we will evaluate the use of AI in the identification of PD biomarkers. An established patient cohort and metabolomic data set is already available for use in this study. Within this cohort plasma and CSF from 159 PD patients and 110 matched controls have been collected longitudinally in 2 year follow-ups over a period of 10 years. Metabolomics data of matched CSF and plasma samples have already been recorded for baseline 2-, 4 and partially 6 year follow up visits and is available for the proposed project. In addition, samples for the 8 and 10 year follow up visits will be available for this project and metabolomics measurements will be contributed by the Hiller group. This combined with the activities of the new PTB/TU Braunschweig planned Junior research group on metabonomic measurement standards will enable results from the preliminary AI assessments to influence a targeted based metabolomics measurement approach, exhibiting a lower measurement uncertainty, which will enable an assessment of the impact on data refinement on the power of AI procedures.

Available data

- Non-targeted MS based metabolomics data of matched plasma and CSF samples for baseline, 2-, 4- and 6-year follow-up samples measured in analytical triplicates. In total ~400 measurements for CSF and ~500 for plasma samples.
- Samples for 8- and 10 year follow-up are available and will be measured upon project start.

Collaboration

- TU Braunschweig BRICS, Prof. Dr. Karsten Hiller,
- Paracelsus-Elena Hospital Kassel, Prof. Dr. Brit Mollenhauer

Candidate Requirements

- PhD in bioinformatics or computing programming,
- Experience in processing of omics data and biochemical measurement procedures would be an advantage.

References

[1] Trezzi, J.-P., Jäger, C., Galozzi, S., Barkovits, K., Marcus, K., Mollenhauer, B. & Hiller, K. Metabolic profiling of body fluids and multivariate data analysis. *MethodsX* **4,** 95–103 (2017).

[2] Trezzi, J.-P., Galozzi, S., Jaeger, C., Barkovits, K., Brockmann, K., Maetzler, W., Berg, D., Marcus, K., Betsou, F., Hiller, K. & Mollenhauer, B. Distinct metabolomic signature in cerebrospinal fluid in early parkinson's disease. *Mov Disord.* **32,** 1401–1408 (2017).

[3] Pietzner M, Engelmann B, Kacprowski T, Golchert J, Dirk A-L, Hammer E, et al. Plasma proteome and metabolome characterization of an experimental human thyrotoxicosis model. BMC Medicine; 2017;15.

[4] Arloth J, Eraslan G, Andlauer TFM, Martins J, Iurato S, Kühnel B, et al.. DeepWAS: Multivariate genotype-phenotype associations by directly integrating regulatory information using deep learning. PLOS Computational Biology. PLOS Computational Biology; 2020;16:e1007616.

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

AI based image enhancement for reduced radiation exposure in CT imaging

Type: Postdoc or PhD

Principal Investigators

- Stefan Pojtinger (PTB, AG 6.25)
- Steffen Ketelhut (PTB, AG 6.25)

Background

Previously, several authors demonstrated that algorithms based on deep neural networks can be used for the enhancement of low-quality medical images [1–4]. The authors proposed certain types of network architectures that can be trained for the reduction of image artifacts and noise in low-dose CT images. For the patient, this means a lower exposure to ionizing radiation, while maintaining the same diagnostic significance of the CT image. The U.S. Food and Drug Administration (FDA) already cleared one such product for PET imaging (SubtlePET, Subtle Medical, USA).

However, algorithms like this carry a certain risk. The enhancement of the image quality with the help of AI based algorithms could suggest a high image quality to the doctor, despite a potential loss of important anatomical detail during the optimization process. E.g. it was shown for AI based algorithms for image reconstruction, that there is "a variety in the failure of recovering structural changes […], ranging from complete removal of details to more subtle distortions and blurring of the features." [5]. To ensure safe diagnostics for the patient, it is essential to explore the limits of such algorithms and to define the necessary requirements.

Project Aim, Objectives and Program

In this project, the suitability of AI based algorithms (in particular deep neural networks) for image optimization in CT imaging will be investigated. Special attention will be paid on the development of test criteria aimed on assessment of robustness of such algorithms (based on methods as described in [5]) which can be used later to increase the confidence in such algorithms and thus accelerate the broad application of AI driven algorithms in hospitals.

The first objective is to identify high-risk algorithms (especially these based on generative adversarial networks, GANs) for image enhancement literature and to establish a database of test-cases. These test-cases do not necessarily include patient data alone, but in addition synthetic test-data obtained from already established Monte Carlo methods and phantom measurements at the departments CT-scanner. As a second objective, the identified algorithms will be implemented and trained on publicly available patient data (cf. available data). The limitations of the algorithms will be examined by application of objective measures for image quality like observer models [6].

Ultimately, the final objective is to derive criteria reflecting the robustness of the implemented algorithms.

Available data

Several databases for medical images are freely available (visceral, LCTSC, Open-Access Medical Image Repositories) and are currently used for training of U-Net based algorithms in our project on automatic image segmentation. In addition, our department maintains contact with various partners such as the German Cancer Research Center (DKFZ).

We provide several facilities like a CT-scanner and several X-ray tubes on site. These facilities can be used for the creation of synthetic data. Also, synthetic data will be generated within the framework of this project using Monte Carlo methods for radiation transport (EGSnrc), that are already established at our department. For this, we provide detailed photon spectra of the source of the departments CT-scanner as well as a ready to use procedure for accessing the corresponding bow-tie filtration [7].

Collaboration

- Heidelberg University, Medical Faculty
- German Cancer Research Center (DKFZ) (current collaborations on AI based segmentation of CT datasets and AI based dose calculation for personalized medicine)
- Technical University of Dortmund, Department of Physics (current collaborations on AI based segmentation of CT datasets and AI based dose calculation for personalized medicine)
- Städtisches Klinikum Braunschweig (collaboration within a research project to unify dose parameters in CT and CBCT)

Candidate Requirements

- MSc. in physics, medical physics or similar
- Affinity to software development
- Basic knowledge in Python, preferably experience with TensorFlow/Keras
- Knowledge of Monte-Carlo software packages like EGSnrc or Geant4 are advantageous

References

[1] Umehara K, Ota J and Ishida T 2018 Application of Super-Resolution Convolutional Neural Network for Enhancing Image Resolution in Chest CT *J. Digit. Imaging* **31** 441–50

[2] Singh R, Wu W, Wang G and Kalra M K 2020 Artificial intelligence in image reconstruction: The change is here *Phys. Medica* **79** 113–25

[3] Zhang Y, Yue N, Su M, Liu B, Ding Y, Zhou Y, Wang H, Kuang Y and Nie K 2020 Improving CBCT Quality to CT Level using Deep-Learning with Generative Adversarial Network *Med. Phys.* mp.14624

[4] Yuan N, Zhou J and Qi J 2020 Half2Half: deep neural network based CT image denoising without independent reference data *Phys. Med. Biol.* **65** 215020

[5] Antun V, Renna F, Poon C, Adcock B and Hansen A C 2020 On instabilities of deep learning in image reconstruction and the potential costs of AI *Proc. Natl. Acad. Sci.* **117** 30088–95

[6] Barrett H H, Yao J, Rolland J P and Myers K J 1993 Model observers for assessment of image quality *Proc. Natl. Acad. Sci. U. S. A.* **90** 9758–65

[7] Rosendahl S, Büermann L, Borowski M, Kortesniemi M, Sundell V M, Kosunen A and Siiskonen T 2019 CT beam dosimetric characterization procedure for personalized dosimetry *Phys. Med. Biol.* **64**

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Deep learning-based dosimetry in medical x-ray imaging

Type: Postdoc or PhD

Principal Investigators

- Stefan Pojtinger (PTB, AG 6.25)
- Steffen Ketelhut (PTB, AG 6.25)

Background

In medical x-ray imaging procedures, it is essential to achieve an acceptable image quality at a minimal dose. The measured application-specific dose quantities (ASD) are dose area product (DAP) in general radiography, fluoroscopy and cone beam computed tomography (CBCT) and dose length product (DLP) in conventional CT. These quantities are conceptionally different and cannot easily be converted into each other.

ASD values are the basis for quality assurance and diagnostic reference levels. However, they are not equivalent to patient dose. Patient dose is usually given in terms of organ or effective dose, which are useful quantities for the determination of potential radiation detriments. Conversion factors are needed to obtain patient dose from measured ASD values. They are usually calculated with Monte Carlo methods for reference x-ray devices and a set of mathematical reference patients. Therefore, they are neither machine nor patient specific and introduce large uncertainties.

An increasing number of imaging tasks can be performed with different x-ray imaging modalities, like CT and C-arm CBCT. The comparison of the radiation risk can often only be achieved using patient dose, as the ASD of different modalities are generally not comparable. Direct methods to determine the individual patient dose are therefore highly desirable. Recently, procedures have been developed which allow quick and precise patient and x-ray machine specific dose estimates in conventional CT [1-4]. These are based on deep learning tools which use the CT image of the individual patient to estimate organ doses by simulation of the machine specific CT scan. Comparable tools are not yet available for other x-ray imaging modalities although there is an increasing demand.

Project Aim, Objectives and Program

The aim of the project is to develop and provide procedures for the quick determination of patient dose for medical x-ray imaging procedures. Quick means that the individualized patient dose in terms of organ and/or effective dose is indicated at the display monitor of the medical x-ray device immediately after the imaging process without causing serious delays in the clinical workflow.

The first objective is to develop methods which correlate 2D or 3D patient images with matching mathematical whole-body patient models. These mathematical models are characterized by pre-segmented organs needed for the evaluation of the organ dose. This task can ideally be solved with tools based on deep learning.

The second objective is to develop procedures for quick 3D dose estimates based on simulations of the imaging process using the pre-determined virtual patient model. This task can ideally be solved with the application of the deep dose estimation algorithm developed by Maier et al [4] which are based on a deep convolutional neural network trained by Monte Carlo dose estimates.

The goal of the third objective is the evaluation of the uncertainty and robustness of the algorithm. This will be done by measurements on anthropomorphic phantoms at real x-ray facilities. Several x-ray facilities are present at PTB, others can be used at the sites of our collaboration partners.

Available data

XCAT provides anthropomorphic voxel models with parametrized anatomical sizes. A reference voxel phantom is provided with the ICRP Report 110.

Several databases for medical images are freely available (visceral, LCTSC, Open-Access Medical Image Repositories) and can be used to develop a method for quick 3D dose estimation. In addition, our department maintains contact with various partners such as the Städtisches Klinikum Braunschweig (SKBS), who could provide anonymized image data of patients.

We have several facilities like a CT-scanner and several X-ray tubes on site, and have numerous different ionization chambers, semiconductor detectors and anthropomorphic and reference phantoms which can be used for validation of the method.

Collaboration

- Heidelberg University, Medical Faculty
- German Cancer Research Center (DKFZ) (current collaborations on AI based segmentation of CT datasets and AI based dose calculation for personalized medicine)
- Technical University of Dortmund, Department of Physics (current collaborations on AI based segmentation of CT datasets and AI based dose calculation for personalized medicine)
- Städtisches Klinikum Braunschweig (collaboration within a research project to unify dose parameters in CT and CBCT)

Candidate Requirements

- MSc in physics, medical physics or similar
- Knowledge of a programming language, preferably Python. Experience with Tensorflow, and Keras are beneficial
- Affinity to software development
- Knowledge of Monte-Carlo based radiation transport codes as EGSnrc or Geant4 are beneficial

References

[1] Z. Peng *et al.*, A method of rapid quantification of patient-specific organ doses for CT using deep-learning-based multi-organ segmentation and GPU-accelerated Monte Carlo dose computing. *Medical Physics*. **47**, 2526–2536 (2020).

[2] S. Sharma *et al.*, A real-time Monte Carlo tool for individualized dose estimations in clinical CT. *Physics in Medicine and Biology*. **64** (2019), doi:10.1088/1361-6560/ab467f.

[3] S. Chen *et al.*, Automatic multi-organ segmentation in dual-energy CT (DECT) with dedicated 3D fully convolutional DECT networks. *Medical Physics*. **47**, 552–562 (2020).

[4] J. Maier, E. Eulig, S. Dorn, S. Sawall, M. Kachelrieß, in *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference, NSS/MIC 2018 - Proceedings* (Institute of Electrical and Electronics Engineers Inc., 2018)

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Project Title: Uncertainty of AI-based dose prediction compared to Monte Carlo methods

Type: Postdoc or PhD

Principal Investigators

- Dr. Hans Rabus (PTB, AG 8.02)
- Prof. Dr. Klaus-Robert Müller (Technische Universität Berlin)

Background

AI-based methods are increasingly used in radiotherapy with a prospect of enabling real-time calculations of dose distributions for adaptive treatment planning [1]. Different from the already more advanced AI-based dosimetry calculations in imaging, treatment planning requires much lower uncertainties on the few percent level. Similar to AI applications in imaging, synthetic reference data sets produced by Monte Carlo simulations will play an important role for training, testing and validating the AI algorithms as systematic discrepancies introduced in data pooling ("center effects", e.g. due to different protocols for data collection and reporting) are absent. In order to simulate at a rate of several thousand reference data sets within a few months, application of special numerical tricks for variance reduction is indispensable that generally require a lot of expertise for their proper application. Many radiation transport codes include respective features that may produce biased results. Such bias on training data as well as the uncertainties of simulation results may introduce another uncertainty component with the predictions of the trained AI algorithms (in addition to those originating in the algorithms themselves). The ground work for establishing uncertainty budgets of radiation transport Monte Carlo simulations has been done in two preceding PhD projects and in code intercomparison exercises in the frame of EURADOS WG 6 Computational Dosimetry [2], [3], [4]. A systematic general framework for establishing uncertainty budgets for Monte Carlo simulations is still to be developed as well as an assessment of the impact of these uncertainties on the predictions of AI algorithms trained by such simulation results.

Project Aim, Objectives and Program

Investigation of the influence of bias and uncertainty in synthetic data sets used for training a neural network on the prediction capabilities of the network for a showcase example from radiotherapy or radiation protection in medicine.

- Investigation of the suitability of different approaches to include uncertainty propagation in Monte Carlo simulations of radiation transport in medical applications.
- Establishment of an uncertainty budget for synthetic reference data sets produced using Monte Carlo simulations.
- Evaluation of the impact of the resulting uncertainties on the uncertainties of AI algorithm-based dose predictions.

Work program:

- Identification of the optimum showcase examples in discussion with EURADOS WGs and (pre-)clinical partners. Examples could be dose distribution in the whole body in a radiotherapy treatment.
- Investigation of possibilities for implementing techniques for uncertainty propagation (e.g. deterministic sampling, sensitivity analysis, …) in Monte Carlo radiation transport simulations
- Development of a detailed Monte Carlo simulation for the showcase. with a radiation transport code implementing the error propagation and variance reduction techniques
- Production of a sufficiently large set of synthetic reference data and training of a neural network or other machine learning algorithm with the set of simulation input data and simulation predictions.
- Assessment of the uncertainties of the reference data and of the resulting uncertainty of the AI-based prediction.
- Investigation of methods for correcting variance-reduction related bias and assessing the uncertainty of the correction.

Available data

- not applicable, they will be produced in the course of the project

Collaboration

- Technische Universität Berlin
- EURADOS WG 6 Computational Dosimetry, WG 9 Radiation dosimetry in radiotherapy, WG 12 Dosimetry in medical imaging and potential new WG on dosimetry in nuclear medicine
- University of Heidelberg and University of Essen
- MIT/Harvard University

Candidate Requirements

- MSc in computer science, mathematics, physics or similar
- Good programming skills and understanding of computer science
- Experience in radiation transport simulations is an asset

References

[1] Vandewinckele L et al. Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance.. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 153, 55–66 (2020)

[2] Villagrasa C et al. Assessing the contribution of cross-sections to the uncertainty of Monte Carlo calculations in micro- and nanodosimetry. Radiation Protection Dosimetry 183, 11-16 (2019)

[3] Caccia B et al. 2020. A model validation scheme for Monte Carlo simulations of a medical linear accelerator: geometrical description and dosimetric data used in the “Linac Action”. EURADOS Report 2020-05. European Radiation Dosimetry Group e. V., Neuherberg.

[4] Rabus H et al. Quality Assurance for the Use of Computational Methods in Dosimetry: Activities of EURADOS Working Group 6 “Computational Dosimetry”. J Radiol Prot in press, doi: 10.1088/1361-6498/abd914 (2021)

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Incorporation of spatial regularization and uncertainty estimations into MR parametric mapping

Type: PhD

Principal Investigators

- Patrick Schünke (PTB, AG 8.13)
- Christoph Kolbitsch (PTB, AG 8.13)

Background

Many quantitative medical imaging techniques rely on a pixel-by-pixel based evaluation of image series. Typical examples include time-dependent signals as in dynamic contrast enhanced (DCE) MRI and dynamic positron emission tomography (PET), frequency-dependent signals as in Chemical Exchange Saturation Transfer (CEST) MRI or more sophisticated dependencies as in magnetic resonance fingerprinting (MRF). Commonly, signal models are fitted to the pixel-based 1D data series or dictionary matching approaches are used to obtain estimates of the underlying quantitative parameters like relaxation times, metabolite concentrations or rate constants.

Recently, neural networks (NN) have been proposed to replace these time-consuming and/or computationally demanding approaches and speed up the parameter estimation[1–3]. Further, the use of NNs allows to obtain uncertainty estimations indicating how confident the NN is about predictions for a certain parameter in a certain pixel[1,3]. Other approaches incorporate spatial regularization into NNs to make use of the high correlation between neighboring pixels and employ these highly valuable information to improve the parameter estimations[4,5].

PTB-8.13 focuses on the development and evaluation of novel approaches to make quantitative MRI (qMRI) more accurate and reliable. In this context, we successfully implemented and tested different AI-based uncertainty estimation approaches for qMRI as well as different regularization approaches for image reconstruction.

Project Aim, Objectives and Program

The goal of this project is to develop a framework for the creation of synthetic multi-dimensional MRI data and their utilization in AI-based MR parametric mapping approaches that incorporate both, spatial regularization, and uncertainty estimations. This requires the following tasks:

- Continued development of a simulation tool for the generation of synthetic multi-dimensional MRI data for supervised learning
- Investigation of different uncertainty estimation approaches and exploration of the universal applicability of these approaches for various MR parametric mapping methods
- Exploration of different regularization methods for neural networks with uncertainty estimation to take spatial correlations into account

Work Program:

- Extension of self-developed simulation tools for MR and MR-PET data[6,7] to simulate various MR signal series (e.g., T_1, T_2, CEST) and creation of virtual MR phantoms with different (anatomical) geometries. Using the extensive know-how of PTB-8.13 about qMRI[8–10] and motion correction[10,11], the option to incorporate realistic noise and artificial motion (artifacts) shall be provided.
- Initial design and training of different Artificial Neural Networks (ANN) with uncertainty estimation using the generated synthetic data and exploration of the universal applicability of the network for various MR parametric mapping tasks.
- Exploration of various approaches like convolutional neural networks (CNN) to incorporate spatial regularization and investigation of their impact on the performance of the parameter quantification and uncertainty estimation in MR parametric mapping.
- Application of the developed and optimized framework to existing and newly acquired multi-dimensional *in vivo* MRI data and investigation of the transferability to other imaging techniques

Available data

- Numerical simulation environment for T_1, T_2, $T_{1\rho}$, CEST, MRF and DCE MRI
- 3D DCE-MRI of the liver (10 patients, 30 pre-clinical data sets)

Collaboration

- Charité – Universitätsmedizin Berlin
- German Cancer Research Center (DKFZ), Heidelberg
- Universitätsklinikum Erlangen

Candidate Requirements

- MSc in physics/computer science or in a comparable field of engineering/natural sciences
- Software experience: Python and joint development (Git)
- Experience in one or more of the fields: machine learning, signal and image processing, MRI

References:

[1] Glang, F. *et al.* DeepCEST 3T: Robust MRI parameter determination and uncertainty quantification with neural networks-application to CEST imaging of the human brain at 3T. *Magn. Reson. Med.* 1–17 (2019) doi:10.1002/mrm.28117.

[2] Shao, J., Ghodrati, V., Nguyen, K. & Hu, P. Fast and accurate calculation of myocardial T1 and T2 values using deep learning Bloch equation simulations (DeepBLESS). *Magn. Reson. Med.* **84**, 2831–2845 (2020).

[3] Bliesener, Y., Acharya, J. & Nayak, K. S. Efficient DCE-MRI Parameter and Uncertainty Estimation Using a Neural Network. *IEEE Trans. Med. Imaging* **39**, 1712–1723 (2020).

[4] Balsiger, F. *et al.* Spatially regularized parametric map reconstruction for fast magnetic resonance fingerprinting. *Med. Image Anal.* **64**, 101741 (2020).

[5] Fang, Z. *et al.* Deep Learning for Fast and Spatially Constrained Tissue Quantification From Highly Accelerated Data in Magnetic Resonance Fingerprinting. *IEEE Trans. Med. Imaging* **38**, 2364–2374 (2019).

[6] pulseq-CEST. https://pulseq-cest.github.io/ (2020).

[7] Mayer, J. *et al.* Flexible numerical simulation framework for dynamic PET-MR data. *Phys. Med. Biol.* **65**, 145003 (2020).

[8] Schuenke, P. *et al.* Adiabatically prepared spin-lock approach for T1ρ-based dynamic glucose enhanced MRI at ultrahigh fields. *Magn. Reson. Med.* **78**, 215–225 (2017).

[9] Schuenke, P. *et al.* Simultaneous mapping of water shift and B1 (WASABI)-Application to field-Inhomogeneity correction of CEST MRI data. *Magn. Reson. Med.* **77**, 571–580 (2017).

[10] Becker, K. M. *et al.* Fast myocardial T1 mapping using cardiac motion correction. *Magn. Reson. Med.* **83**, 438–451 (2020).

[11] Ippoliti, M. *et al.* 3D nonrigid motion correction for quantitative assessment of hepatic lesions in DCE-MRI. *Magn. Reson. Med.* **82**, 1753–1766 (2019).

Physikalisch-Technische Bundesanstalt
Braunschweig und Berlin
Nationales Metrologieinstitut

Project Title: Accelerating radiation transport simulations in radiation medicine by machine learning

Type: PhD

Principal Investigator

- Dr. Hans Rabus (PTB, SeSc 8.02)

Background

Training, testing and validation of AI methods require large data sets. In particular with regard to AI methods for therapy planning and real-time radiation dosimetry in interventional radiology, measured data on patients is only available to a limited extent for data protection and ethical reasons. In addition, the field of in vivo dosimetry is still in its infancy and in principle only allows point measurements. Synthetic reference data sets obtained using Monte Carlo simulations of radiation transport will therefore be essential for AI procedures in radiation medicine in the future. As was highlighted in several presentations at the recent ESTRO congress (online, Nov 28th -Dec 1st, 2020), synthetic reference data are advantageous, as systematic discrepancies introduced by different protocols for data collection and reporting ("center effects") are absent. Furthermore, they can, in principle, be produced continuously and at a higher rate than data derived from patients.

While modern Monte Carlo radiation transport codes or toolkits offer the possibility to perform very detailed simulation, the method faces the major drawback that such detailed simulations are very CPU time intensive [1], [2]. The major challenge to be solved is therefore to accelerate the Monte Carlo simulations to quickly generate new data sets in sufficient numbers (for example, if data sets previously used for validation of AI methods have become known).

Project Aim, Objectives and Program

Aim: To enable real-time detailed Monte Carlo simulations.

Objective: Accelerate Monte Carlo simulations by coupling them with machine learning techniques such that intermediate results of the simulation are used to train a neural network. Demonstrating the proof-of-principle for a representative showcase example (to be identified in interaction with clinical partners)

Short description of work program:

- Investigation of possible interfaces between Monte-Carlo (MC) codes for radiation transport simulation and machine learning (ML) methods
- Investigation of the suitability of neural networks for coupling with MC simulations
- Identification of a radiation transport code suitable for coupling with ML methods
- Development of a MC simulation that interacts with a neural network
- Proof-of-principle for a concrete application, e.g. dose calculation in flexible anthropomorphic numerical phantoms

Available data

- not applicable, they will be produced during the project

Collaboration

- Technische Universität Berlin (Prof. Dr. Klaus-Robert Müller)
- German Cancer Research Center (Biophysics in Particle Therapy, Dr. Andrea Mairani)

Candidate Requirements

- MSc in physics (preferential), mathematics, computer science
- Good background in physics
- Experience in one or more of the fields: Radiation transport simulations, Neural networks

References

[1] Javaid U et al. Mitigating inherent noise in Monte Carlo dose distributions using dilated U-Net. Medical Physics 46, 5790–5798 (2019)

[2] Neishabouri A et al. Long short-term memory networks for proton dose calculation in highly heterogeneous tissues. Medical Physics in press, doi: 10.1002/mp.14658 (2021)